Doctoral theses at NTNU, 2023:356

Dipendra Jee Mandal

# Image Quality Assessment of Hyperspectral and Conventional Imaging for Cultural Heritage Artifacts

Doctoral thesis

**NTNU**
Norwegian University of Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of
Science and Technology

Dipendra Jee Mandal

# Image Quality Assessment of Hyperspectral and Conventional Imaging for Cultural Heritage Artifacts

Thesis for the Degree of Philosophiae Doctor

Gjøvik, November 2023

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of
Science and Technology

**NTNU**
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Information Technology and Electrical Engineering
Department of Computer Science



ISBN 978-82-326-7412-1 (printed ver.)
ISBN 978-82-326-7411-4 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)

Doctoral theses at NTNU, 2023:356

Printed by NTNU Grafisk senter

*In loving memory of my dear mother,* ***Madhuri Devi Mandal****,*
*whose enduring love and support forever surround and inspire me.*

# Abstract

Preserving historical artifacts is crucial as they carry cultural significance beyond their artistic value and offer insight into their creation for future generations. Digitalization of cultural heritage artifacts preserves them by enabling high-resolution digital images for enhanced accessibility, broader dissemination, and improved conservation. It also offers adaptability for various applications and reduces the need for physical handling, minimizing the risks of damage and long-term deterioration. However, the digital representation of cultural heritage faces challenges in the adaptation of technology, economic sustainability, standardization, and quality control. To ensure accuracy, completeness, fidelity, and legibility compared to the original object, integrating quality assessment is essential during the digitization life cycle of cultural heritage. Therefore, this thesis seeks to investigate and identify essential parameters for characterizing and evaluating the behavior and content of digitized artifacts in the domain of cultural heritage. The research presented in this thesis covers both RGB imaging and spectral imaging techniques, focusing on two important cultural heritage artifacts; paintings and microfiche.

For artwork such as paintings, it is essential to identify and classify pigments accurately. It helps curators and conservators determine the historical context, authenticity, and proper conservation methods. Conventional RGB imaging systems are valuable in various applications, including cultural heritage. Nevertheless, their limited spectral range makes them impractical when additional information is needed, such as material identifications, composition analysis, detection of hidden features, etc. Spectral imaging, such as multispectral and hyperspectral imaging, overcomes these limitations by capturing much broader spectral information. In this thesis, hyperspectral imaging is used to analyze paintings. The spectral quality of the imaging system was evaluated on its accuracy in identifying pigments for classification purposes. Various spectral metrics based on supervised algorithms and machine learning models were used and analyzed for pigment classification.

Likewise, despite having a long lifespan of 500+ years, microfiche materials are susceptible to physical degradation caused by light exposure, temperature fluctuations, improper handling, and poor storage conditions. Consequently, digitization becomes necessary. Microfiche relies on specialized devices for direct human eye reading, primarily available in select archives or libraries. However, over time, these devices can suffer from various problems that affect their usability. Given the impor-

tance of preserving and accessing microfiche materials, in this thesis, we investigated alternative options for microfiche digitization. The process of digitizing microfiche involves magnification, which can introduce various factors that can impact the overall reproduction quality, such as noise, distortion, and artifacts. Additionally, these materials can suffer from poor legibility and clarity because of their reduced size, potential degradation over time, or even limitations of the scanning device. To address this, objective and subjective image quality assessments are considered in this thesis to assess the quality of the digitized microfiche.

The objective of this Ph.D. research is to evaluate the quality of digitization of cultural heritage artifacts. Through a compilation of articles, this dissertation contributes to overall research objectives and seeks to provide a comprehensive understanding of the research subject to assess the quality of digitized cultural heritage artifacts using conventional and hyperspectral imaging technologies. The analysis and practical recommendations of these articles discussed in this thesis are coherent and provide valuable resources for researchers and practitioners involved in the preservation and documentation of cultural heritage and those involved in imaging, and provide valuable insights into potential directions for future research.

# Sammendrag

Bevaring av historiske artefakter er avgjørende ettersom de har kulturell betydning utover deres kunstneriske verdi og gir innsikt i deres skapelse for kommende generasjoner. Digitalisering av kulturarvartefakter bevarer dem ved å muliggjøre høyoppløselige digitale bilder for bedre tilgjengelighet, bredere spredning og forbedret bevaring. Det tilbyr også tilpasningsevne for ulike bruksområder og reduserer behovet for fysisk håndtering, og dermed minimerer risikoen for skade og langvarig forringelse. Imidlertid står den digitalisering av kulturarv overfor utfordringer med teknologi, økonomisk bærekraft, standardisering og kvalitetskontroll. For å sikre nøyaktighet, fullstendighet, trofasthet og lesbarhet sammenlignet med det originale objektet, er det avgjørende å integrere kvalitetsvurdering som en viktig fase i digitaliseringssyklusen til kulturarv. Derfor fokuserer denne avhandlingen på å undersøke og identifisere essensielle parametere for å karakterisere og evaluere atferden og innholdet til digitaliserte artefakter innenfor kulturarvområdet. Forskningen presentert i denne avhandlingen dekker både RGB-bildebehandling og spektralbildebehandling, med fokus på to viktige kulturarvartefakter, nemlig malerier og mikrofiche.

For kunstverk som malerier er det viktig å identifisere og klassifisere pigmenter nøyaktig. Dette hjelper kuratorer og konservatorer med å bestemme historisk kontekst, ekthet og riktig bevaringsmetode. Konvensjonelle RGB bildebehandlingssystemer er verdifulle for mange områder, inkludert kulturarv. Likevel er deres begrensede spektrale rekkevidde upraktiske når ytterligere informasjon er nødvendig, som materialeidentifikasjon, sammensetningsanalyse, deteksjon av skjulte informasjon osv. Spektral bildebehandling, som multispektral og hyperspektral bildebehandling, overvinner disse begrensningene ved å fange bredere spektral informasjon. I denne avhandlingen brukes hyperspektral bildebehandling for å analysere malerier. Spektral kvalitet i bildebehandlingssystemet ble evaluert med tanke på nøyaktighet i identifikasjon av pigmenter for klassifisering. Forskjellige spektrale metrikker basert på veiledede algoritmer og maskinlæringsmodeller ble brukt og analysert for pigmentklassifisering.

På samme måte, til tross av å ha lengre levetid, er mikrofiche-materiale sårbart for fysisk nedbrytning forårsaket av lyspåvirkning, temperatursvingninger, feil håndtering og dårlige lagringsforhold. Derfor er digitalisering nødvendig. Mikrofiche er avhengig av spesialiserte lesere for direkte avlesning av menneskeøyne, primært tilgjengelig i utvalgte arkiver eller biblioteker. Imidlertid kan disse enhetene over tid oppleve

ulike problemer som påvirker brukervennligheten. Gitt viktigheten av å bevare og få tilgang til mikrofiche-materiale, undersøkte denne avhandlingen alternative måter for digitalisering av mikrofiche. Prosessen med å digitalisere mikrofiche innebærer forstørrelse, noe som kan introdusere ulike faktorer som kan påvirke den overordnede reproduksjonskvaliteten, som støy, forvrengning og artefakter. I tillegg kan disse materialene lide av dårlig lesbarhet og klarhet på grunn av redusert størrelse, potensiell nedbrytning over tid eller til og med begrensninger i skanneutstyret. For å håndtere dette, gjøres objektive og subjektive bildekvalitetsvurderinger i denne avhandlingen for å vurdere kvaliteten på digitaliserte mikrofiche.

Dette doktorgradsarbeidet er en samling artikler som bidrar til de overordnede forskningsmålene og søker å gi en omfattende forståelse av forskningsobjektet for å vurdere kvaliteten på digitaliserte kunstverk ved bruk av konvensjonelle og hyperspektrale bildebehandlingsteknologier. Analysen og de praktiske anbefalingene av artiklene som diskuteres i denne avhandlingen er sammenhengende og gir verdifulle ressurser for forskere og praktikere som er involvert i bevaring og dokumentasjon av kulturarv, samt de som er involvert i bildebehandling, og gir verdifulle innsikter i potensielle retninger for fremtidig forskning.

# Acknowledgements

Completing this Ph.D. journey has been a long and challenging endeavor, and I am thankful to everyone who has supported me along the way. This research would not have been possible without their guidance, encouragement, and assistance.

First and foremost, I would like to thank my main supervisor, Prof. Marius Pedersen, from NTNU, Norway, for his invaluable guidance and support throughout this academic journey. I will forever be grateful for his encouragement and belief in my ability, which motivated me to move forward during my difficult situations and finally reach this milestone. I would also like to sincerely thank my co-supervisors, Prof. Sony George from NTNU, Norway, and Dr. Clotilde Boust from C2RMF, France, for their unwavering support, guidance, and deep insights.

I thank Dr. Hilda Deborah for her exceptional mentorship throughout this PhD journey. The hours you spent advising me, your constructive feedback, and your assistance in navigating the intricacies of academic research have been helpful. I am also grateful to my dissertation committee members: Prof. Markku Hauta-Kasari from the University of Eastern Finland (Finland), Ass. Prof. Susan Farnand from the Rochester Institute of Technology (USA) and Ass. Prof. Andras Ladai from NTNU (Norway) for their valuable feedback throughout the process of evaluating my dissertation.

I am sincerely grateful to the European Union for funding my research through the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 813789 as part of the CHANGE project. I also thank IDI, NTNU for their support in funding my doctoral research. Their belief in the importance of this research work is greatly appreciated.

I extend my thanks to the CHANGE management team, including Prof. Jon Yngve Hardeberg, Prof. Sony George, Anneli Torsbakken Østlien and Anne Hilde Ruen Nymoen, their lovable and smiling faces, filled with warmth and support, made the journey even more enjoyable and encouraging. I am also grateful to NTNU, especially Colourlab, for providing an enriching academic environment and resources that greatly facilitated my research. Additionally, I would like to acknowledge the exceptional contributions of the administrative and library staff and the IT support teams, who played a crucial role in ensuring a seamless research process.

I extend my thanks to my colleagues and fellow researchers at Colourlab and the fellow ESRs from the CHANGE and ApPEARS projects, who have been a source of inspiration and camaraderie. The discussions, debates, and shared experiences that we have had together have not only enriched my research, but also contributed significantly to my personal growth. Thank you for being like a family to me.

My heartfelt gratitude goes to my parents, Madhuri Devi Mandal and Ram Prakash Mandal; their enduring support and unwavering faith in my abilities have not only been a source of motivation, but also a reminder that with determination and perseverance, I can overcome any obstacle that comes my way. Finally, I thank my siblings, especially my sisters Mintu and Ruby, for their love and support, as well as all of my other family members and friends.

Thank you all for being part of this remarkable journey.

Gjøvik, Norway
October 2023
Dipendra Mandal

# Contents

# Figures

# Tables

# Acronyms

**CH** Cultural Heritage. xv, 1–5, 11–13, 16, 17, 20, 21, 26, 29, 30, 35, 36, 38, 41, 47, 48, 50, 51, 53–55, 58–62

**CHANGE** Cultural Heritage Analysis for New Generations. 1

**CRISATEL** Conservation Restoration Innovation System for imaging capture and digital Archiving to enhance Training Education and lifelong Learning. 22

**FADGI** Federal Agencies Digitization Guidelines Initiative. xv, 30

**GRSS** Geoscience and Remote Sensing Society. 31

**HSI** Hyperspectral imaging. xi, xiii, 4, 5, 7, 11, 13, 21–24, 30, 31, 35, 36, 38, 47, 49, 50, 57–63

**IEEE** Institute of Electrical and Electronics Engineers. 31

**IQM** Image Quality Assessment. 3, 11, 27

**ISO** International Organization for Standardization. xv, 3, 30

**NIST** National Institute of Standards and Technology. 31

**VASARI** Visual Arts System for Archiving and Retrieval of Images. 22

**VNIR** Visible and Near-Infrared. 4, 7, 22, 35, 47, 57, 58

# Part I

# Overview

# Chapter 1

# Introduction

This Ph.D. research is part of the European Union-funded Marie Skodowska-Curie Actions, Innovative Training Network project entitled Cultural Heritage Analysis for New Generations (CHANGE). The project is coordinated by the Norwegian University of Science and Technology (NTNU), within the Department of Computer Science (IDI), Gjøvik, with the participation of the Center for Research and Restoration of Museums of France (C2RMF), France, and other beneficiaries throughout Europe. This thesis addresses a specific objective within the CHANGE project, which is to assess the quality of Cultural Heritage (CH) artifact digitization.

This dissertation is a compilation of articles that together contribute to the overall research goal and aim to provide a comprehensive understanding of the research topic. The analysis and practical recommendations of these articles, as cohesive, serve as a valuable resource for researchers and practitioners involved in preserving and documenting CH artifacts, and provide insights into possible directions for future research. This chapter provides an overview of the research work carried out during the Ph.D. and is divided into five sections: Motivation (Section 1.1), Research Objectives (Section 1.2), Research Framework (Section 1.4), List of Publications (Section 1.3), and Thesis Organization (Section 1.5).

## 1.1 Motivation

In essence, digitization means the process of converting analog signals or data into digital signals or data (Lee 2001). In the context of CH, digitization refers to creating digital surrogates of physical materials, such as documents, photographs, and other artifacts. CH digitization is not recent; museums, libraries, and archives have created digital versions of their collections for several decades. The practice of digitizing CH materials began in the 1960s, when institutions started exploring the use of computer systems to manage and preserve their valuable resources. Electronic catalogs were an early form of digital documents created in the 1970s and continued

with the conversion of printed source material into digital files in the 1980s (Horik 2005). Before the advent of digital technologies in the 1960s, microforms [1] were one of the available methods for archiving and preserving large documents. In the 1930s, the Library of Congress microfilmed several books and manuscripts, making it one of the earliest known applications of microfilm for the preservation of cultural heritage (Brown et al. 2012). For a comprehensive overview of digital imaging and its application in the CH domain, refer to the books [2] (Kenney et al. 2000; MacDonald 2006).

Historical artifacts, such as sculptures, paintings, documents, and manuscripts, carry great cultural significance beyond their artistic value, and serve as vital resources to understand the cultural context in which they were created. However, these works of art are vulnerable to degradation caused by human handling, natural aging, and environmental conditions. Thus, safeguarding these artifacts is of utmost importance in preserving our CH. One of the effective solutions is to use various digital imaging techniques to digitize these CH artifacts, which can play a crucial role in their preservation, documentation, and utilization. By creating high-resolution digital images, digitization enables improved accessibility, wider dissemination, and improved conservation, thereby safeguarding these artifacts for future generations. An effective approach to mitigate these artifacts' risks is using various digital imaging techniques to digitize them. This process aids in their preservation, documentation, and utilization. Additionally, the use of digital imaging techniques reduces the need to handle artifacts physically, particularly those that are fragile and delicate, further mitigating the risk of damage and deterioration over time, furthermore providing a means to protect and disseminate their cultural significance to forthcoming generations.

The lack of space and inadequate documentation pose challenges to accessibility, resulting in a considerable portion of museum collections being kept in storage. Consequently, only a few museums can display a significant part of their entire collections. To improve the accessibility of stored collections, museums have implemented various strategies, including visible storage, lending, and exchange of items (Corona 2023). Although these efforts contribute to addressing the issue, digitization of collections offers an essential solution by creating virtual exhibitions that make these artifacts accessible to a broader audience without the limitations of physical space. Despite the inherent advantages, the digital representation of CH artifacts presents several challenges. These include adapting to modern imaging technology, addressing economic considerations related to maintaining large digital repositories, establishing standardized digitization procedures, and implementing effective quality control

[1] Microforms are media formats, such as microfilm, microfiche, microcards, and microprints, that are used to store and preserving information in compact and reduced formats.

[2] These books provide a comprehensive overview of digital imaging for libraries and archives, covering topics such as user requirements, digital camera principles, high-resolution imaging, image processing limits, image databases, color management, image compression, conservation imaging, online monument accessibility, panoramic visualization, virtual restoration, and research policy. They provide extensive information and guidance on these topics and their theoretical constructs, guiding the reader in their practical application while being informed of previous practice and theory.

measures (Conway 2011). One crucial aspect of quality control is evaluating the accuracy of the digital version compared to the original object, with completeness, fidelity, and legibility being critical factors (ISO19264 2021). Therefore, it is imperative to introduce quality assessment as an essential stage in the life cycle of CH digitization.

Conventional RGB imaging [3] system has proven useful in many applications in CH (Derrien 1993; Martinez 1991). However, its usefulness may be limited when more information is required, for example, material identification, addressing metamerism, revealing underdrawings, and pigment classification [4], which cannot be achieved with the limited spectral range of RGB imaging. Spectral imaging systems, such as multispectral and hyperspectral, can overcome this limitation. Image Quality Assessment (IQM) has been an active research area for several decades, showing significant progress, and continues to develop in many fields (Z. Wang and A. C. Bovik 2006). However, defining image quality remains a complex subject. In general, the concept of quality is subjective and context-dependent. The International Organization for Standardization (ISO) definition reflects this by defining quality as the complete set of traits and attributes of a product or service that are relevant to its capacity to meet explicitly or implicitly expressed needs. In other words, a product is considered to have good quality if it meets the requirements specified by the customer (ISO9000 2015). In terms of an image, quality can be assessed subjectively and objectively, which can be further assessed in the spatial or spectral domain. Although the quality criteria for RGB imaging are often subjective, reflecting the visual perception of a human observer (Engeldrum 2004; Z. Wang, A. Bovik, et al. 2004), this is not the case for spectral imaging, as it captures data beyond the visible range and is used for a wide range of applications (Khan et al. 2018), which makes it challenging to generalize the definition of quality. In the literature, various definitions of spectral quality can be found (Shrestha, Pillay, et al. 2014; Stefanou et al. 2009; Sweet et al. 2000). For instance, one such definition delineates quality as the suitability of a specific dataset for a particular purpose (Fryskowska et al. 2018), which is more appropriate for spectral imaging in general.

Preserving and restoring paintings are critical endeavors that come with various challenges. Such as the removal of dirt and old varnish without damaging the paint layer (Stulik et al. 2004) and the selection of appropriate materials for retouching (Digney-Peer et al. 2020). To address these problems, it is essential to accurately identify the pigment used by the artist in an artwork. Scientific analysis sometimes requires physical samples; however, due to the nature of CH objects, it is not recommended to take samples from the object, which in fact destroys the object even at a microscale, and so very often it is not permitted (Quye et al. 2019; Tite 2002).

[3] In this thesis, conventional RGB imaging refers to capturing and displaying images using three additive primary colors; red (620-750 nm) green (495-570 nm), and blue (450-495 nm).

[4] In the context of this thesis, pigment classification refers to the identification of the specific pigment used in the artwork by analyzing its spectral characteristics and then accurately grouping the pixels within the artwork according to the correct pigment class to which they belong.

Therefore, HSI techniques have gained significant attraction due to their non-contact and non-destructive imaging capability, and several studies have demonstrated the effectiveness of HSI in CH applications (Bai et al. 2017; George et al. 2019; Picollo et al. 2020).

Microfiche, a type of microform (De Haas 1958), was widely used to archive and preserve historical documents before the introduction of advanced digital technologies. Given the fragile state of these objects, gaining access to numerous historical documents stored in libraries around the world can be challenging, and in some cases, manuscripts or fragments have even been lost (Garcia-Spitz 2017; Tigchelaar 2004). As a result, microfiche copies of these documents may be the only surviving records of these objects or the sole accessible option for a broader audience, rendering them valuable artifacts and essential components of our cultural heritage. Although microfiche has a long lifespan, it is still susceptible to damage (NEDCC 2017), and requires digitization to ensure preservation and broader accessibility. The need to find solutions to these issues motivates the research and thus this thesis aims to address some of the identified issues by examining and quantifying potential problems. It proposes practical workflows, tools, and methodologies to mitigate them effectively, ultimately contributing to the preservation and accessibility of CH.

## 1.2 Aims and Research Questions

The aim of this Ph.D. research is to investigate and identify crucial parameters for characterizing and evaluating the behavior and content of digitized artifacts in the CH domain. The primary objective of the research is to evaluate the quality of digitization of CH artifacts, which is further divided into two main areas. The first area is the evaluation of image quality in the Visible and Near-Infrared (VNIR), and the second is the assessment of image quality in the visible region. These areas encompass four main research questions.

**RQ1 How does acquisition parameters impact the accuracy of pigment classification in digitizing artworks using HSI ?**

The heterogeneity of pigment structures in artwork and the use of varnish layers can cause specular reflection, especially due to the angle of illuminations (Kubik 2007), affecting the spectral accuracy and resulting in incorrect pigment classification. Objects at varying depths from the camera may appear out of focus in an imaging device with a low depth of field if they are away from the focus plane (Ray 2011; Webb et al. 2020). Capturing CH objects requires maintaining a sharp focus across the entire object. However, variations in image sharpness can occur due to irregularities of objects, resulting in blurry images and degraded quality (Martinez and Hamber 1989). Furthermore, objects in CH applications are sensitive to temperature, and prolonged exposure to illu-

minants during acquisition can cause material property changes and damage. Therefore, investigating how acquisition parameters affect spectral data and classification accuracy is crucial in CH applications.

**RQ2 What is the effectiveness of existing classification algorithms for pigment classification in artwork, and how can they be optimized ?**

Many supervised-based classification algorithms exist for HSI, mostly in remote sensing applications, for example, mineral identification (Melgani et al. 2004; Tripathi et al. 2019). However, some of these algorithms are being adopted directly or with some modification in other application domains such as medical imaging (N. Liu et al. 2020; Zhi et al. 2007), food and agriculture (Melit Devassy et al. 2020; B. Park et al. 2007), and forensics (Deepthi et al. 2022). HSI acquisition in CH is done in a controlled environment with close proximity to the object, allowing for controlled illumination. In contrast, remote sensing HSI is collected under natural illumination with a greater distance from the target, causing atmospheric effects and temporal illumination variations. The classification algorithms used in remote sensing may not be directly applicable or effective in CH due to these differences. For example, an algorithm insensitive to intensity variation may perform well in remote sensing but not in CH. Because of faded or aged pigments (Weerd et al. 2005), pure pigment mixed with different binding mediums (Cosentino 2014a), mixed pigments, that is, pigment mixed with percentages of different weights of lead white (Cavaleri et al. 2013), etc. can have variations in magnitude, which is essential for both diagnostic and conservation purposes. Few of these algorithms have been used for the identification of pigments in artwork using HSI, and therefore it is necessary to explore and evaluate them.

**RQ3 How does the elevation affect the accuracy of pigment classification in artworks ?**

CH artworks are not limited to two-dimensional canvases or boards, which implies that they may not always be flat. The addition of relief, which creates 2.5D or 3D effects, is a significant consideration (Rubens n.d.; Townsend 1995). Various factors can contribute to this third dimension (Barrett et al. 1995; Baxter et al. 2004; Sinclair 1995), such as the morphological textures of brushstrokes on the painted surface (Fu et al. 2018) and the application of a thick layer of pigments to create depth by artists (impasto technique) (Elkhuizen et al. 2019; Gonzalez et al. 2019; Groen 1997; Plisson et al. 2014). The geometry of a relief created through brush painting, impasto techniques, or other factors can affect the interaction of light with the surface, which consequently might influence the spectral signature captured by the hyperspectral sensor for a given pixel. Given the importance of pigment identification or classification in the artwork, it is essential to investigate the impact of such factors. Additionally,

this research question is also motivated by the feedback received during the publication review process (Article A) within this thesis.

**RQ4 What are the available imaging technologies for microfiche digitization and how can they be improved to improve the quality of the resulting images ?**

The content documented in a microfiche is too small to be read by the naked eye due to its significantly reduced size (Saffady et al. 1978), and requires a special device to enlarge, scan, and print the microforms in readable format. Access and availability to microfiche readers or other similar microform reader machines are limited today, although they were once a central aspect of archiving. Such devices are designed primarily for reading rather than digitization, and only a handful of specialized archives or libraries may still have microfiche readers to access their microform collections. Therefore, it is necessary to explore other imaging devices as a feasible solution to digitize microfiche and enhance accessibility to historical collections in addition to traditional microfiche readers. Image enhancement through image processing is a well-established technique that is used to improve visual quality. This is achieved by adjusting various image attributes, such as brightness, contrast, and sharpness, which can produce a more aesthetically pleasing result for a given scenario (Singh et al. 2014). Given the importance of digitizing microfiche collections for preservation and accessibility purposes, exploring the potential of image enhancement techniques in this domain can be highly beneficial.

## 1.3 List of Publications

This research study resulted in five main articles that advocated the objective established in Section 1.2 and constitute the core of this thesis. Three of those articles have been published, one has been accepted for publication, and one is under revision. The summary and analysis of these articles can be found in Chapter 3 of this thesis. In addition, a supporting article is listed that is not part of the core thesis, but is closely associated with the work presented here. Finally, a book chapter is submitted for publication that summarizes the entire thesis.

**Core Contributing Articles:**

Article A: D. J. Mandal, S. George, M. Pedersen, and C. Boust (2021). “Influence of Acquisition Parameters on Pigment Classification using Hyperspectral Imaging.” In: *Journal of Imaging Science and Technology* 65.5, 050406-1–050406–13

Article B: D. J. Mandal, M. Pedersen, S. George, H. Deborah, and C. Boust (2023). "An Experiment–based Comparative Analysis of Pigment Classification Algorithms using Hyperspectral Imaging." In: *Journal of Imaging Science and Technology* 67.3, pp. 030403-1–030403-18. DOI: 10.2352/J.ImagingSci.Technol.2023.67.3.030403

Article C: D. J. Mandal, M. Pedersen, S. George, and C. Boust (2023). "Comparison of Pigment Classification Algorithms on Non–Flat Surfaces using Hyperspectral Imaging." In: Manuscript under revision in Journal

Article D: H. Deborah and D. J. Mandal (2021). "Evaluation of Text Legibility in Alternative Imaging Approaches to Microfiche Digitization." In: *Proc. IS&T Archiving*, pp. 96–101. DOI: 10.2352/issn.2168-3204.2021.1.0.22

Article E: D. J. Mandal, H. Deborah, and M. Pedersen (2023). "Subjective Quality Evaluation of Alternative Imaging Techniques for Microfiche Digitization." In: *Journal of Cultural Heritage*. Accepted for publication

**Supplementary Contributions:**

Article : A. Kadyrova, M. Pedersen, B. Ahmad, D. J. Mandal, M. Nguyen, and P. H. Zimmermann (2022). "Image enhancement dataset for evaluation of image quality metrics." In: *Electronic Imaging* 34, pp. 1–6

Book Chapter : D. J. Mandal, S. George, M. Pedersen, and C. Boust (2023). "Quality Evaluation in Cultural Heritage Digitization." In: *Cultural Heritage Analysis for new Generations (CHANGE)*. Submitted for publication. Routledge

## 1.4 Research Framework

Figure 1.1 illustrates an overview of the research structure and the publication originating from a specific topic of interest. Quality assessment of cultural heritage artifacts is carried out in both VNIR and visible regions, specifically on painting mockups and microfiches, respectively.

For VNIR, the acquisition was made using HSI in a laboratory setup for flat and elevated mockups. The first step involved applying a classification algorithm for pigment classification and analyzing the impact of various acquisition parameters on pigment classification accuracy, which resulted in Article A and addressing research question RQ1. Different supervised and machine learning-based models were then applied to compute pigment classification and tested on the flat region of a newly prepared mockup B with selective pigments, resulting in Article B and addressing the research question RQ2. A similar procedure was carried out for elevated regions and the results were published in Article C, addressing the research question RQ3.

In addition, we digitized the microfiche materials containing typewritten text in the visible region using various imaging devices. Objective image quality assessment was

performed using metric OCR, and the results were published in Article D. However, since microfiche materials can also contain handwritten text and photographs, our study included microfiche with natural scenes and ancient handwritten fragments in addition to typewritten text. A subjective quality assessment was conducted using pair comparison methods to address the limitation of off-the-shelf OCR systems. The results are incorporated into Article E. Both Article D and Article E addressed the research question RQ4.

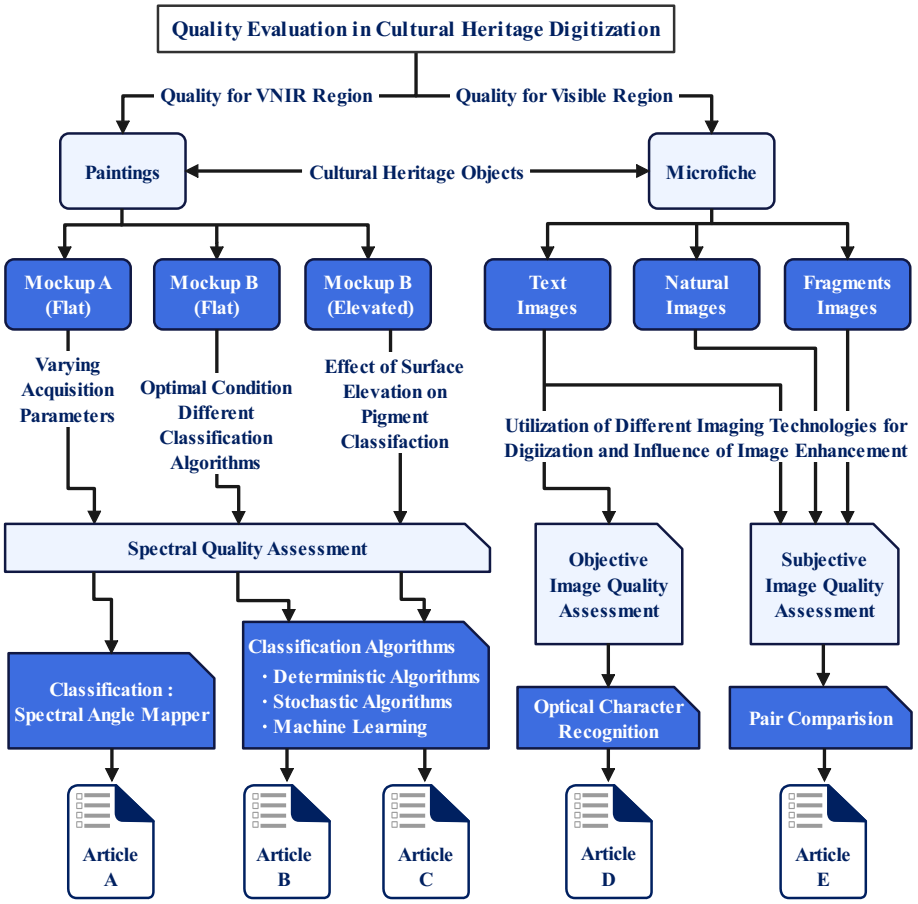



**Figure 1.1:** A diagrammatic representation of the research structure addressed in this thesis and their relationship to the articles included.

## 1.5 Thesis Organization

This thesis is organized into two main parts. Part I comprises five chapters, with Chapter 2 providing essential background knowledge to aid the reader in understand-

ing the research work. Chapter 3 presents a summary of the articles included in the study, while Chapter 4 discusses the contributions made to the field and the research results obtained. Lastly, Chapter 5 contains the study's conclusions, which include a summary of the key findings and their implications and recommendations for future research. Part II comprises all the core contributing articles and is included as an attachment.

# Chapter 2

# Background

In order to ensure a clear understanding of the presented research work, it is crucial to establish a comprehensive understanding of the fundamental concepts and technologies underlying this field. Therefore, this chapter provides an introduction to these topics, facilitating the reading and understanding of the subsequent content presented in the thesis.

In this chapter, we first present the Cultural Heritage (CH) artifacts used in this study. After that, we provide a brief overview of the different imaging technologies used for the acquisition of these artifacts. We begin by explaining the principles of electromagnetic radiation and its role in imaging, followed by an introduction to conventional three-channel and Hyperspectral imaging (HSI), highlighting their differences, advantages, limitations, and application in the context of CH studies. In addition, the chapter briefly discusses the preprocessing of hyperspectral data, which is a critical step in the data processing pipeline. We then provide a brief overview of Image Quality Assessment (IQM), covering objective and subjective quality assessment and various image enhancement techniques. Finally, the chapter provides an overview of classification algorithms and their accuracy assessment methods.

## 2.1 Cultural Heritage Artifacts

CH refers to the collection of cultural resources and traditions that have been inherited from previous generations. It serves as a representation of our past, present, and the continuation into the future, forming a vital link between these temporal dimensions (Blake 2000). The concept encompasses both tangible and intangible forms of a society's cultural expressions that hold historical, cultural, aesthetic, scientific, and social importance. CH plays a crucial role in shaping a society's identity and influencing its values, norms, and customs (Prott et al. 1992). It is widely recognized as a communal asset that belongs to all members of a community, and therefore

we must all conserve and safeguard it for future generations (Jeffers 2015). CH is often associated with artifacts, such as paintings, sculptures, historical monuments, and archaeological sites. However, this understanding of CH fails to capture its full breadth and depth. Over time, the definition of CH has expanded to include all forms of evidence of human creativity and expression. This includes photographs, documents, books, manuscripts, and instruments as singular objects and comprehensive collections. In this thesis, we focus mainly on studying and analyzing two distinct valuable CH elements: paintings and microfiche.

### 2.1.1 Painting

Tangible artwork, such as paintings, has significant artistic value and offers unique insights into the social, cultural, and historical context of their creation (Fichner-Rathus 2008). Paintings serve as historical repositories, influenced by human actions and natural forces. Analyzing how time and the elements have shaped these artworks provides valuable insight into the properties of materials and artistic techniques used during their creation (Taft et al. 2000). A painting consists of several materials, such as the support (canvas, wood, etc.), ground or primer, pigment (providing color), binder (adhering pigment to the support), vehicle (liquid medium used to mix with pigment and binder), varnish (protecting the surface), and mediums for altering paint properties. Brushes and palette knives are used for the application, while an easel supports the artist to work comfortably. The artist's skill and these materials combine to shape the final artwork's appearance and expression. In paintings, while other materials are undoubtedly essential, pigments, on the other hand, play a crucial role as integral components of any artwork. Therefore, their analysis is of utmost importance.

#### Pigments in Artworks

The use of pigments in artwork dates back thousands of years (Feller 1986; Harley 1982), and over time various natural and synthetic pigments have been developed and used. Pigments are a vital component of such works of art, as they provide color and contribute to its overall aesthetic appeal (Feller 1986). In the past, most of the powdered pigments used were derived from minerals, plants, and animals sources and subsequently mixed with organic compounds derived from animals or plants as a binding agent (Barnett et al. 2006). Generally, pure pigments were dominated by a single mineral compound, which possessed distinctive absorption coefficients and absorption bands (Cosentino 2014b). Combining various minerals was a common practice among artists to produce a diverse range of pigment shades, resulting in distinct spectral reflectance curves.

The study of pigments in the artwork can provide valuable insights into the history and development of art and the cultural and social context in which it was created. For example, some pigments were more expensive or difficult to obtain, indicating the social and economic status of the artist or patron. Pigments are also sometimes

used for symbolic or religious purposes, reflecting the beliefs and values of the culture that produced the artwork. Analyzing the composition and properties of pigments can help researchers determine the age, authenticity, and origin. It can also help conservators develop appropriate preservation strategies to protect and extend the useful life of these important cultural artifacts. Pigment mockups are physical replicas of a painting, created with materials and methods that mimic the ones used in the original painting (Stoveland, Stols-Witlox, et al. 2021). To avoid any risk of damage to the original painting, mockups are often used instead of real paintings. Conservators, scientists, and researchers have more flexibility to try different materials, methods, and approaches on the mockup, which can help to identify the most effective approach before moving on to the actual painting (Stoveland, Frøysaker, et al. 2021). As an illustration, Figure 2.1 shows one of the pigment mockups used in this research work.


**Figure 2.1:** Pigment mockup B; mockup was prepared using pigment tubes composed of high-stability pigments and oil, purchased from Zecchi (*ZECCHI* 2023). The pigments were selected based on three main factors: their prominence in CH research articles, their specific spectral characteristics, and consultation with experts in the CH field. Veridian(P1), Cerulean Blue(P2), Green Earth(P3), Yellow Ochre Light(P4), Burnt Umber(P5), Ultramarine Blue Deep(P6), Lead White Hue (P7), Genuine Vermilion(P8), Cobalt Blue Deep(P9), and Ivory Black(P10) are pigments used in the mockup. The linen canvas was primed using three layers of white gesso before applying these pigments.

### Pigment Classification

Pigment classification refers to the categorization and identification of pigments based on their spectral characteristics. Essentially, the objective is to take an HSI image (datacube) and turn it into a meaningful map, where each pixel in the map is part of some class (distribution of the pigments within the works of art). Each class has a unique spectral signature that represents different pigment types. The user can

define these classes or automatically be defined based on spectrally similar pixels (for example, in terms of spectral reflectance). Figure 2.2 illustrates an example of pixel-based classification.

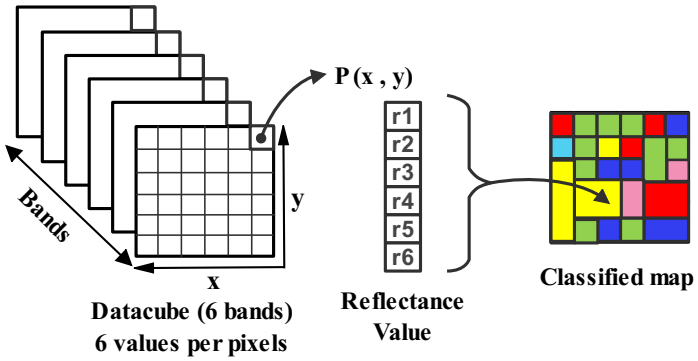



**Figure 2.2:** An illustration of the classification for a 6-band datacube, where each pixel is characterized by six spectral bands corresponding to its reflectance values. Using these distinct values, classes are assigned to the pixels, which are then mapped onto a raster image.

In paintings, we commonly come across both pure and mixed pigments. Analyzing pure pigments involves a direct comparison of their reflectance spectrum with a known spectral signature library. However, when mixed pigments are used, the measured reflectance spectrum does not simply result from the addition of the reflectance of two distinct materials based on their concentrations. Instead, it follows a nonlinear relationship due to scattering effects, necessitating various linear and nonlinear spectral unmixing models (Chen et al. 2019; Lyu et al. 2020; Rohani et al. 2018). To effectively apply these models, a considerable amount of prior knowledge is needed about mixtures, such as information about particle size, mass, and molecular structure (Gueli et al. 2017). In this thesis work, only pure pigments are used in the mockup to simplify the analysis.

Classification can be divided into two groups based on how classes are defined: supervised and unsupervised classification. In supervised classification, the user provides reference or training data for each class and runs algorithms to perform the classification. In contrast, unsupervised classification automatically identifies classes solely based on the spectral properties of the pixels without any provided reference or training data. For this thesis, only supervised classification was used. For a supervised pigment classification, the first step involves defining classes based on the number of pigments present. Subsequently, collect the spectra for each class to create reference spectra, also known as endmembers or spectral libraries. These endmembers are typically derived from the known molecular structure of pigments (Goetze et al. 2022) or by capturing point reflectance over a set of wavelengths within a region of interest from various paintings. For this research, end members were directly

extracted from a specific region of interest within a mockup. Once the end members are obtained, algorithms assign each pixel of the acquired datacube to its respective predefined class.

#### Classification Algorithms

Spectral similarity plays a crucial role in various domains, including remote sensing, spectroscopy, and image analysis, particularly in the context of classification algorithms. Its purpose is to quantify the degree of similarity between the spectral responses of different samples or pixels within a spectral dataset. For classification, several algorithms or metrics (Jagalingam et al. 2015; Shrestha, Pillay, et al. 2014) have been developed to compute spectral similarity. These algorithms differ in their capacity to account for shape and magnitude variations between the two spectra being compared. The approaches used for matching spectra can be classified as deterministic, i.e., based on geometrical and physical aspects, or stochastic, based on distributions (Vishnu et al. 2013). Among the algorithms that are most frequently used are the Spectral Angle Mapper (SAM) (Kruse et al. 1993), Spectral Correlation Mapper (SCM) (De Carvalho et al. 2000), and the Spectral Information Divergence (SID) (Chein-I Chang 2000), among others. Due to the significant achievement of artificial intelligence models in classification tasks, machine learning, and neural network-based approaches have recently been applied to the classification task in works of art (L. Liu et al. 2023). Support Vector Machine (SVM) (L. Wang 2005), Neural Network (NN) (L. Liu et al. 2023), and Convolutional Neural Network (CNN) (Hu et al. 2015; Yu et al. 2017) are a few of the common techniques used.

### 2.1.2 Microfiche

Microform refers to a broad range of formats that contain microreproductions of documents and images. These formats can be made of different materials, including plastic film (such as microfilm and microfiche), paper (such as microcard and microprint), or a combination of both (such as an aperture card or a microfilm cell embedded in a punchcard) (Auger 1991; Spigai 1973). Microfiche is a type of microform consisting of a flat film sheet that is typically 4 x 6 inches in size (which is according to the American National Standard (IT9.2 1988)); however, in Europe, the size may vary from 3 x 5 inches (which is very common) up to 8 x 6 inches. The sheet contains numerous microimages arranged in a grid pattern, which can be photographs, drawings, or documents that have been scaled down in size, often by a factor of 24×, but up to 48× as well, to enable the storage of a significant amount of information in a small space (British Library 1992; De Haas 1958). As an illustration, Figure 2.3 shows one of the microfiche used in this research work.

Microfiche has been in existence for a long time, with its earliest form being invented around the late 18th century. Nevertheless, it was not made commercially available until the mid-1930s, and it was not until the 1950s that it started to be commonly utilized. The microfiche (or microfilm) comprises different layers, with the base layer

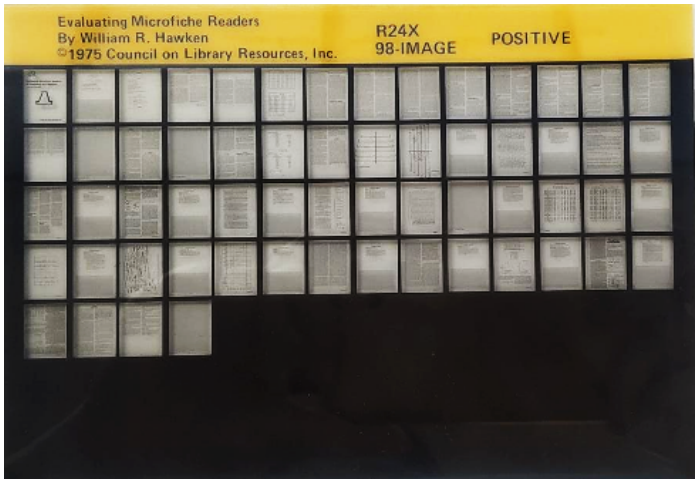



**Figure 2.3:** An example of a positive microfiche, composed of microimages of all pages of the handbook (Hawken 1975b);measuring 105mm × 148mm in physical size and featuring a reduction ratio of 24X.

and the emulsion layer being the most important. The base layer is the thickest and provides stability for the film, while also serving as a foundation for the other layers. These layers are typically made of acetate (used from the 1930s to the 1980s) or polyester (used from the 1960s to the present). The emulsion layer, on the other hand, is responsible for holding the actual image and is chemically fixed to the base layer during the development process. Although most microfiche uses silver-gelatin emulsions, there are other types, such as diazo microfiche, which uses dye-based and light-sensitive materials, and vesicular microfiche, which are thermally processed and heat-sensitive (Archives 2013; Selle 2003). Microfiche remains available today; however, there has been a consistent decrease in the production of microforms overall since the 1990s due to the advent of digital technology. Before then, microforms were the primary means of archiving and preserving large documents, such as newspapers. The CH domain quickly adopted it to capture its collections for preservation, ease of access, and wider distribution. As a result, many historical collections can now only be accessed in microfiche (Garcia-Spitz 2017).

### Microfiche Reader

A microfiche reader is necessary to read the content or pages of a microfiche because of its significant reduction ratio. The reduction ratio is a measure of the linear relationship between the original document size, and its reduced format or microimage (Hawken 1975b). Typically, microfiche readers consist of a light source, optics, and a projection mechanism that projects the image onto a viewing screen. To view the text and images on the microfiche, the user can adjust the focus and magnification of optics; at the same time, this can be a time-consuming process. Some

microfiche readers have the capability to print or save images for further analysis or reference. The Zeutschel delta plus microform reader (Zeutschel 2023a), which was used in this work and is available through the local library in Gjøvik, Norway is shown in Figure 2.4. This device is marketed for digitizing all formats of microfilms and photographic materials, with the microfiche listed as a compatible input type. It can support a reduction ratio of 7× to 105×.


**Figure 2.4:** The delta+ Advanced model of the film scanning system designed for various microforms. The delta+ Advanced model is an upgraded version of the delta plus. The study, however, utilized the delta plus model of the system. *Image Source* (Zeutschel 2023b)

## 2.2 Imaging

Imaging techniques are used to digitize or document CH artifacts. The imaging process involves the interaction of electromagnetic radiation with the object or scene being examined. This interaction causes the object to reflect, absorb, transmit and emit incident radiation in varying proportions, depending on the radiation's characteristics and the object's properties (Gauglitz et al. 2006). A digital imaging sensor captures a portion of the incoming radiation and converts it into an electrical signal, which is then digitized. In the context of preserving CH, various forms of electromagnetic radiation and imaging techniques can be utilized to create a digital representation of an object's external appearance and internal structure (Verhoeven 2016). This further helps to gain a deeper understanding of the composition and structure of CH materials, as well as other objects.

### 2.2.1 Electromagnetic Radiation

Electromagnetic radiation refers to the form of energy composed of oscillating electric and magnetic fields that travels through space in the form of waves or particles at the speed of light(c). James Clerk Maxwell, a Scottish physicist who formulated the theory of electromagnetism in the 19th century, explains a significant characteristic of these waves, stating that the electric field's direction is perpendicular to the magnetic field that generates it and vice versa. Furthermore, the direction of propagation of electromagnetic waves is perpendicular to electric and magnetic fields (forming a transverse wave), and they oscillate in phase (Fitzpatrick 2008). A visual representation of an electromagnetic wave is shown in Figure 2.5.

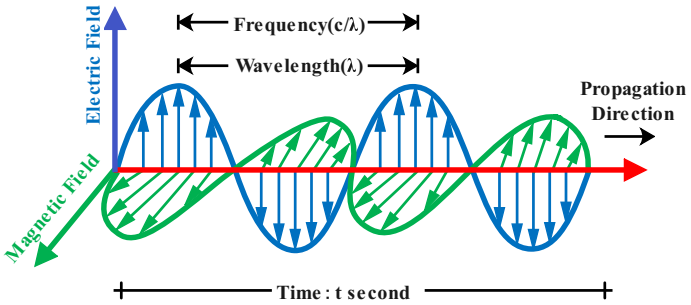



**Figure 2.5:** Electromagnetic radiation consists of changing electric and magnetic fields; the oscillating electric fields (blue) and magnetic fields(green) are orthogonal and propagate as an electromagnetic wave in the direction indicated by the arrow, travelling at the speed of light.

Electromagnetic waves can be characterized by either the wavelength ($\lambda$) or the frequency ($f$) of their oscillations to form the electromagnetic spectrum, ranging from very long wavelengths, such as radio waves, to very short wavelengths, such as gamma rays. $f$ and $\lambda$ are related to each other through the universal wave equation (applies to all waves) and is given by Equation (2.1). Here $v$ is the velocity (speed of a wave); however, for an electromagnetic wave, we use $c$ (speed of light) to represent the speed instead of $v$.

$$v = \lambda \times f \tag{2.1}$$

Figure 2.6 illustrates the electromagnetic spectrum, categorized according to their respective wavelengths or frequencies; this includes radio waves, microwaves, infrared, visible light, ultraviolet, X-rays and gamma rays. Each segment of the electromagnetic spectrum has distinct properties and interactions with matter, leading to various applications in diverse fields such as telecommunications and imaging. The boundaries between these adjacent segments of the electromagnetic spectrum are not precisely defined or universally agreed upon, so the ranges tend to overlap.

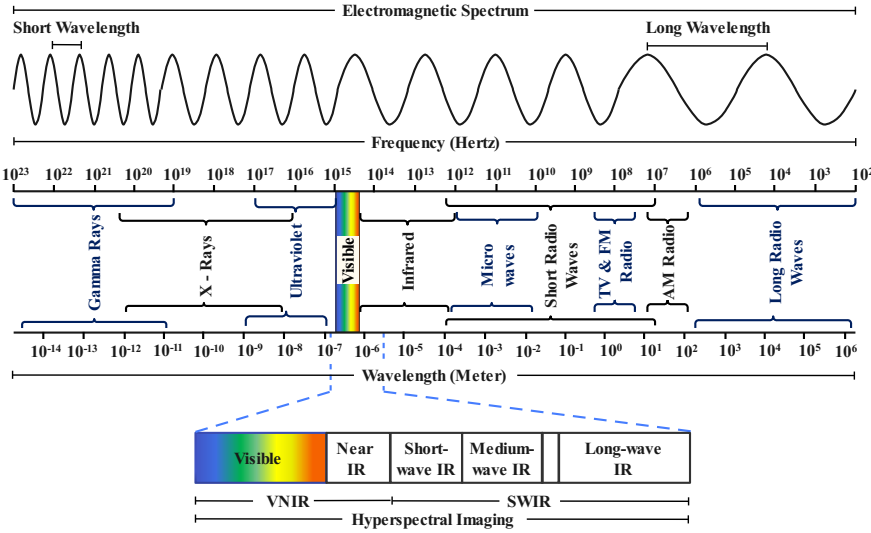



**Figure 2.6:** The electromagnetic spectrum showing segments of the spectrum with associated wavelengths and frequencies. Each segment represents a specific portion of the spectrum, including radio waves, microwaves, infrared, visible light, ultraviolet, X-rays, and gamma rays. Spectrum highlighting the VNIR (Visible and Near-Infrared) and SWIR (Short-Wave Infrared) regions holds significance in studying and preserving cultural heritage using hyperspectral imaging. *Image adapted from* (DiGiuseppe 2011)

Figure 2.7 illustrates the various extents of penetration of the electromagnetic spectrum into a layered structure of an easel painting (Schreiner, Frühmann, et al. 2004). The layers depicted in the figure encompass a comprehensive range of potential layers that can exist within a painting's structure, serving further as a representative model of the overall structure of a painting. An adhesive layer, usually composed of glue, is applied to the uppermost portion of the canvas, serving two primary purposes: to line the painting or to establish the ground layer. The application of a ground layer serves various functions, such as safeguarding the underlying support materials from oil diffusion in oil painting. Following this, a sequence of sketches, called underdrawings, is created on top of the ground layer. These sketches are commonly made using carbon-rich materials such as charcoal or graphite. An initial translucent glaze called Imprimitura is applied to prepare the painting for subsequent layers, which acts as a primer agent. Paint layers typically consist of a mixture of powdered pigment and a binder, selected according to the specific painting medium used, whether acrylic, oil, or watercolor. Artists opt for binders that best suit their chosen medium, including oils, synthetic binders, or gum arabic. Finally, a layer of varnish is added to the painting, serving multiple purposes, such as protecting the paint layers from dirt and dust accumulation and acting as a protective shield for the artwork (Baias 2020; Gilroy et al. 2017). For a comprehensive exploration of the application of different ranges of electromagnetic spectrum utilized by different

imaging techniques in the analysis of paintings, please refer to (MacBeth 2012).

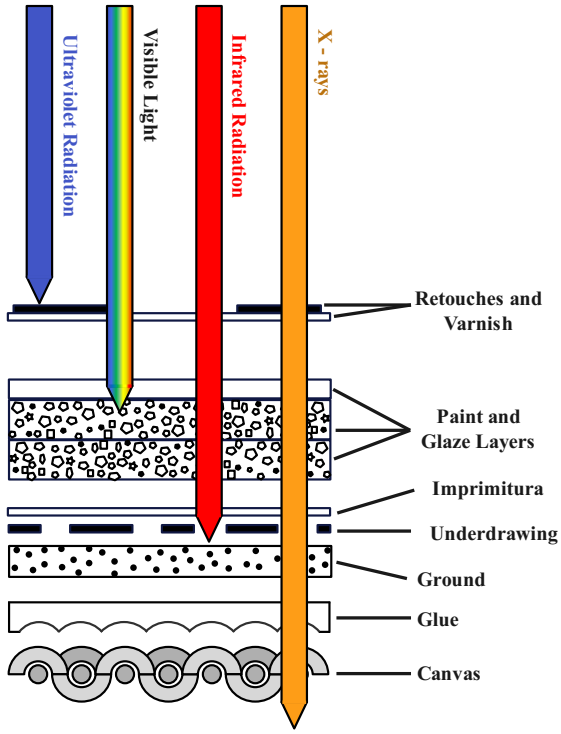



**Figure 2.7:** Structure of an easel painting depicting different layers and its penetration of electromagnetic radiation at different wave lengths. UV radiation is commonly utilized for rapid diagnostics of varnish layers; radiation within the visible region, known as light, enables the perception of colors and spatial details. IR radiation exhibits an ability to penetrate deeper into the layers of the painting, facilitating examination of the underlying structure, composition, and hidden layers. X-radiation has higher penetrating capabilities, allowing for diagnostics of diverse artworks, such as supports, statues, monuments, and more. *Image adapted from* (Schreiner, Frühmann, et al. 2004; Schreiner, Wiesinger, et al. 2017)

### 2.2.2 Imaging Techniques

In the context of CH imaging, the goal is often to generate a digital record of an object that faithfully captures its visual features, including color, texture, and hidden features such as material identification, pigment analysis, underdrawings, and more, without damaging or altering the object being imaged. The selection of imaging techniques in CH depends on the specific goals of the user and how well these techniques align with the unique properties of the object being documented. Spectrophotometers are widely used in CH to measure the absorption and transmission of light by an object

across a range of wavelengths. They offer several nanometer resolutions and a wide range of wavelengths from ultraviolet (UV) to infrared (Boust et al. 2023). Although UV, infrared, and RGB imaging can be valuable tools for studying paintings and artwork, HSI in the VNIR (Visible and Near-Infrared) or SWIR (Shortwave Infrared) regions offers several distinct advantages (Hayem-Ghez et al. 2015). For example, different pigments and materials can exhibit similar behavior in UV, infrared, or RGB images, making them difficult to distinguish using these techniques. Similarly, some pigments may have overlapping absorption bands, which can create challenges in distinguishing them using lower-spectral-resolution data. Compared to imaging systems that capture data in a few discrete wavelength bands, HSI provides higher spectral resolution, which means that data are collected in numerous narrow and closely spaced bands across the spectrum. This increased spectral resolution allows for finer differentiation between spectral features, making it easier to distinguish between closely related pigments with subtle differences in their spectral signatures. In the following section, a brief description of RGB and spectral imaging is provided.

#### Conventional RGB Imaging

The conventional imaging technique used in CH, also called RGB imaging or visible-light photography, relies on the electromagnetic spectrum within the visible region. This region encompasses wavelengths approximately ranging from 380 to 700 nanometers (nm). During the process, an image sensor or digital camera captures light and divides it into its red (at around 670nm), green (at around 550nm), and blue (at around 448nm) wavelength components. This separation is accomplished by using either a filter array or three distinct image sensors, where each sensor is dedicated to capturing one specific color channel. Subsequently, the intensities of these three components are combined to produce a color image. For a detailed exploration of the concepts and applications of RGB imaging, please refer to (Reinhard et al. 2008).

Although conventional RGB imaging plays a crucial role in capturing, analyzing, and presenting color information in various CH artifacts, however, due to its limited spectral range, which is composed of only three channels, it can become impractical in certain cases (Linhares et al. 2020; Yamaguchi et al. 2008). For example, this limitation becomes evident in cases where distinguishing between pigments of a similar hue becomes necessary. Similarly, metamerism poses another challenge, as it causes identical colors to appear different under different lighting conditions. Furthermore, when users require more comprehensive information beyond color, such as material identifications, underdrawings, etc., this imaging may not be suitable. Spectral imaging systems such as multispectral and hsi can be used to overcome these limitations.

### Spectral Imaging

Spectral imaging systems extend the spectral range (i.e., it is not limited to the visual spectrum, and can also be used in near infrared, infrared and ultraviolet spectrum of electromagnetic radiation) and enable a more comprehensive analysis and interpretation of CH artifacts. Multispectral imaging can be achieved through different approaches (Hardeberg et al. 2002; Shrestha and Hardeberg 2013). One method involves using traditional optical filters in a filter wheel or using a tunable filter (Hardeberg 2001; Hardeberg et al. 2002), which is placed in front of a high-resolution digital camera or a stereo camera (Hardeberg et al. 2002; Shrestha, Mansouri, et al. 2011). Another promising technique utilizes multiplexed light-emitting diode (LED) illumination (J.-I. Park et al. 2007). In this method, a series of LEDs at distinct wavelengths are sequentially illuminated, while a monochrome camera captures images under each illuminated LED. This process results in the generation of a multispectral image with $n$ number of bands.

In order to achieve a high level of color fidelity in artworks, specifically paintings, the Visual Arts System for Archiving and Retrieval of Images (VASARI), project funded by the European Union (1989-1992), has developed a multispectral system utilizing a monochrome digital camera paired with seven filters across the visible range (Martinez, Cupitt, et al. 2002). Likewise, another EU-funded project called Conservation Restoration Innovation System for imaging capture and digital Archiving to enhance Training Education and lifelong Learning (CRISATEL),(2001-2005), has developed a multispectral setup featuring a monochrome digital camera equipped with 13 interference filters (Liang et al. 2004).

### Hyperspectral Imaging

Hyperspectral imaging (HSI), also referred to as imaging spectroscopy, is a non-invasive imaging technique that combines the features of imaging and spectroscopy to acquire spatial and spectral information from an object and generates a spatial map over continuous and narrow spectral bands, producing a three-dimensional datacube or hypercube, that is, $X\ (x, y, \lambda)$, $x$ and $y$ represent spatial information as rows and columns (or vice versa), while $\lambda$ represents the spectral dimension (Qin 2010b). The wavelength range of a hyperspectral device is determined by the specific sensor it uses and the intended spectral range to which it is designed to capture. Certain hyperspectral devices are specifically designed for Visible and Near-Infrared (VNIR) imaging, which typically spans approximately 400 to 1000 nanometers (nm) (vnir 2023). Alternatively, other device may be tailored for shortwave infrared (SWIR) imaging, covering a broader range from around 1000 to 2500 nm (swir 2023). Furthermore, there are hyperspectral devices available that can capture imagery in the mid-wave infrared (MWIR)(SPECIM 2023a) and long-wave infrared (LWIR) (SPECIM 2023b) ranges, extending beyond 2500 nm and reaching several micrometers.

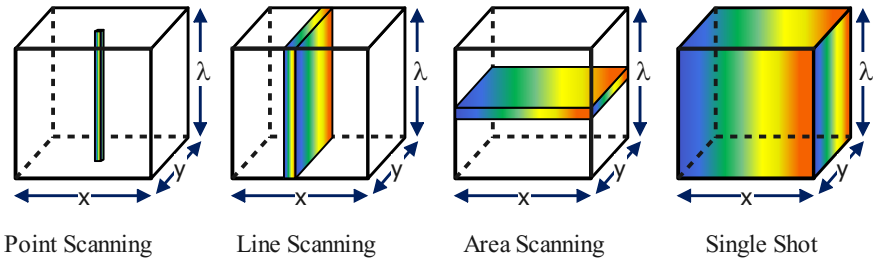



**Figure 2.8:** Four distinct methods for obtaining a HSI datacube; the spatial dimensions are denoted by x and y, while the $\lambda$ denotes the spectral direction, representing the wavelength. *Image adapted from* (Qin 2010b)

As illustrated in Figure 2.8, the HSI datacube can be acquired in four different ways, i.e., point scanning (whiskbroom), line scanning (pushbroom), area scanning (band sequential), and single shot methods. Point scanning involves obtaining a spectrum of a single point in an object; the process is repeated for each point in an object, and the object is moved in the x and y directions point-by-point using a computer-controlled stage. Line scanning involves acquiring spectral measurements from a single line of an object at a time that are simultaneously recorded by an array detector, and either object or a device is moved line-by-line. In area scanning, an image of the entire object is captured using a two-dimensional detector array for each band, without having to move the object. The process is repeated for several spectral bands, resulting in a series of two-dimensional images combined to form a hyperspectral datacube. The single-shot method is a rapid acquisition technique that captures a HSI in a single exposure. It utilizes a two-dimensional array detector that simultaneously captures the entire object in multiple spectral bands (Qin 2010b). Each method offers distinct advantages and trade-offs with respect to speed, spatial resolution, spectral resolution, and cost.

Multispectral and hyperspectral imaging can be distinguished in terms of the number of bands captured and the spectral resolution. Multispectral imaging captures a limited number of broad spectral bands (usually more than three and up to 20), while hyperspectral imaging captures a larger number of narrower and contiguous spectral bands (up to a few hundred), providing much higher spectral resolution. Figure 2.9 illustrates the comparison between RGB, multispectral, and hyperspectral imaging, emphasizing the differences in terms of the number of bands used and the continuity of the spectral waveform captured by each method.

### 2.2.3 Hyperspectral Data Preprocessing

Before hyperspectral data can be utilized for subsequent data analysis, they need to undergo preprocessing. This process comprises several steps in which the obtained datacubes with digital number (DN) are transformed into normalized reflectance values. Figure 2.10 illustrates the general processing steps that have been adopted in the

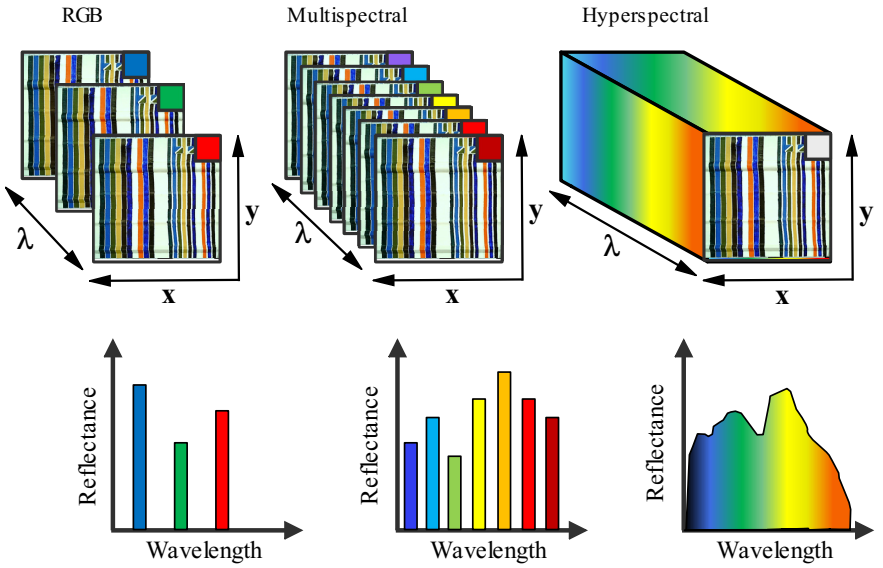



**Figure 2.9:** Comparison of conventional RGB, multispectral and hyperspectral imaging. RGB imaging captures visible light using three color channels (red, green, and blue), while multispectral imaging utilizes multiple discrete spectral bands, including those beyond the visible spectrum. HSI achieves higher spectral resolution by capturing images in numerous narrow and contiguous bands, enabling detailed spectral analysis. *Image adapted from* (Q. Li et al. 2013)

course of this thesis work. This preprocessing is essential to ensure the suitability of the hyperspectral data for further analysis and to facilitate accurate and reliable information extraction. In this section, we consider the data preprocessing primarily linked to the pushbroom mode of data acquisition. Periodically, device-level calibrations are performed, often by manufacturers, to ensure accurate performance of hyperspectral sensors. One such important calibration is the spectral calibration, which addresses any distortions in the sensor's output that depend on the wavelength (Qin 2010a). This calibration procedure involves the use of reference calibration sources or materials with known spectral signatures that are compared to the measured spectra obtained from the sensor. Various setups can be employed for spectral calibration, such as directing narrowband sources like Hg gas discharge lamps or lasers with well-defined spectral lines into an integrating sphere. Through this process, adjustments can be made to correct spectral distortions such as spectral shift, band-to-band misalignment, or spectral smile (Baumgartner et al. 2012).

The dark current, caused by the spontaneous generation of electrons within the detector chip as a result of the thermal effect, leads to an electrical signal within the sensor even in the absence of light. This phenomenon increases with longer integration times and higher operating temperatures of the device (Widenhorn et

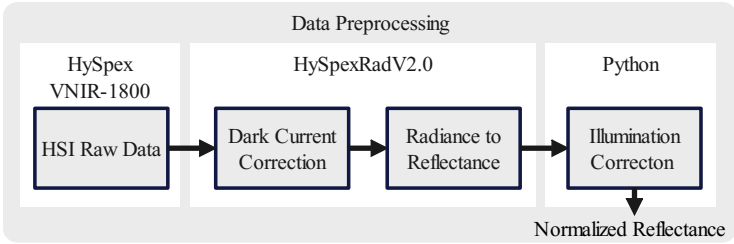



**Figure 2.10:** Hyperspectral data preprocessing steps for the retrieval of normalized reflectance.

al. 2010). To account for this, the dark current needs to be subtracted from each pixel in subsequent scan data. In the case of the VNIR1800 (vnir 2023) device used in this research, the acquisition software has a built-in feature for automating and simplifying this step. It involves closing an automatic shutter and capturing an average of 200 frames without any light illuminating the sensor, which is then saved to a file. Reflectance factor, also known as relative reflectance, is a measure of the proportion of reflected light intensity from a surface compared to the reflected light intensity from a calibration reference at different wavelengths. It quantifies the amount of incident light that is reflected, irrespective of factors such as illumination conditions, sensor sensitivity, or target distance. To normalize the acquired data to the reflectance factor, a known reference, such as calibrated reflectance standards, must be utilized. Calibration targets commonly employed, such as Spectralon®(Spectralonl 2023) or Zenith Polymer®(SphereOptics 2023) reflectance standards, consist of matte Lambertian reflecting surfaces. These materials ensure that the reflected light has nearly equal intensity in all directions.

Hyperspectral data obtained from airborne or satellite platforms can be affected by atmospheric scattering and absorption, and to mitigate these atmospheric effects, atmospheric correction algorithms (Kale et al. 2017) are used. However, for close-range applications, such as acquisition of paintings, the influence of the atmosphere is not as significant, and therefore atmospheric corrections are not necessary. On the other hand, geometric calibration (Špiclin et al. 2010) is essential to ensure accurate spatial alignment of the hyperspectral data. It involves correcting distortions caused by sensor motion and platform orientation. Some hyperspectral camera manufacturers provide acquisition software that automatically calculates the scan speed, which is also the case with the device we used. However, not all manufacturers offer this feature, which requires users to manually input the scan speed. Furthermore, incorporating a geometric calibration target within the scene during the scanning process can aid in correcting any potential geometric errors during the subsequent postprocessing stage. For more detailed information on the calibration workflow specifically for artwork acquisition using hyperspectral imaging, please refer to (Pillay, Hardeberg, et al. 2019).

## 2.3 Understanding Quality

The term quality is frequently used in various aspects of our daily lives, but it is often not clearly defined. In a study (R. W. Hoyer et al. 2001), the authors examine the viewpoints of several prominent contributors to the advancement of quality in the twentieth century. The study proposes two broad definitions of quality. Firstly, quality can be understood as the production of products or the delivery of services with measurable characteristics that comply with a predetermined set of specifications, often expressed numerically. Second, quality can be defined as meeting customer expectations about using or consuming a product or service. Regardless of the specific field, the commonly accepted definition of quality is "conformance to requirements" (R. W. Hoyer et al. 2001). It is important that the requirements are explicitly stated to avoid any misunderstanding. We can effectively assess and measure quality by specifying requirements in terms of numerical specifications.

### 2.3.1 Image Quality

In broad terms, image quality can be defined in terms of visibility of distortions, or artifacts, such as color shifts, blurriness, nosiness, blockiness, and geometry, etc., introduced in an image when it is captured, processed, stored, compressed, transmitted, displayed, and printed. Valuable insights and a diverse contribution from scholars and experts in different fields have led to numerous interpretations and definitions of image quality within the literature (Keelan 2002; Pedersen 2011; Phillips et al. 2018). Depending upon the particular use case and the context in which it is applied, the most appropriate definition of image quality can be formulated and/or adopted.

(Janssen 2001) described image quality in the context of the visual-cognitive system, where visual information of an image received through the visual system is processed by our brain in terms of two attributes, i.e., usefulness and naturalness. The usefulness of an image corresponds to the precision of the internal representation of the image, and the naturalness of an image is the degree of correspondence between the internal representation of the image and knowledge of reality as stored in memory. Image quality, as defined by (Keelan 2002) and adopted by the ISO 20462 (ISO20462 2005) photography standard, is an overall impression of the merit or excellence of an image. It is based on the perception of an observer who is not directly associated with the process of capturing the image (photography) or closely involved with the subject matter depicted in the image. Involving an independent observer helps to eliminate any potential biases or familiarity that could influence the judgment.

Image quality in the context of CH refers to the degree of fidelity and effectiveness in the digital representation of cultural artifacts such as paintings, manuscripts, photographs, or other significant historical objects. Here, degree of fidelity refers to how well they reproduce the original records in terms of accuracy for attributes such as tones, color, details, etc. Effectiveness, on the other hand, signifies the importance of these attributes with respect to specific use cases. For instance, when digitizing

paintings created with watercolors, achieving a high degree of accuracy in capturing color information is typically crucial. However, this level of importance may not be equally applicable when digitizing newspapers or documents, where the legibility of the captured text is of greater significance.

### 2.3.2 Spectral Image Quality

As mentioned in Section 2.3.1, quality criteria for conventional RGB imaging often rely on subjective measures reflecting human visual perception, which may not apply directly to the spectral domain. For spectral imaging, the meaning of quality varies depending on the specific application, indicating the need for distinct considerations when assessing image quality for different purposes. This challenge becomes particularly evident in the case of hyperspectral imaging, which captures data beyond the visible range and finds applications in various fields.

For instance, a researcher aiming to identify signs of a widespread disease in crops may prioritize the spectral resolution and spectral band of the imagery rather than its spatial resolution, as it is crucial for detecting the impending stress (Mishra et al. 2017). Conversely, a military analyst searching for small objects may place greater emphasis on spatial resolution rather than spectral content (Yan et al. 2021). Thus, an image with numerous bands but moderate spatial resolution might be considered as a high quality for one analyzing spectral content. However, an image with limited bands but high spatial resolution could be considered as even higher quality for an analyst examining spatial information (Kerekes et al. 2004). Consequently, establishing a universally applicable definition of spectral image quality becomes a complex task.

According to (Fryskowska et al. 2018), quality is defined as the extent to which a specific dataset is suitable for a particular purpose. This definition can be applied to spectral imaging, highlighting the importance of image or data utility for task-oriented objectives. For example, when performing pigment classification in an artwork using hyperspectral imaging, spectral data are considered of high quality if they yield a high classification accuracy.

## 2.4 Image Quality Assessment

Image Quality Assessment (IQM) has been an active area of research for several decades, showing substantial progress and ongoing development. Quality assessment is crucial to validate technological advances over system accuracy. Image quality can be evaluated subjectively and objectively, with further assessment in either the spatial or spectral domains. The human visual system (HVS) (Wandell 1995) is the ultimate receiver of visual signals in most visual communication systems, and therefore the most basic method of assessing the quality of an image is to present it to an expert human observer. However, human perception can vary among and to address this issue, one can overcome subjectivity by gathering multiple perspectives

from different individuals and statistically analyzing the outcomes. This approach is known as a subjective image quality assessment. It involves presenting the image to a group of individuals with/without expertise in image evaluation, who then provide their subjective opinions or ratings regarding the image quality. To ensure accuracy and consistency in the assessment, statistical analysis techniques are applied to the collected data.

Several methods exist in the literature to measure image quality (Keelan 2002; Mantiuk et al. 2012; Pedersen 2011; Pinson et al. 2003), such as paired comparison, rank ordering, categorical sort, and magnitude estimation. In this thesis work, we used a force-choice paired comparison (David 1988) method without ties. This method involves displaying two images side by side to the participants and asking them to compare the images and select the one they perceive as having high quality or being more visually appealing. If the participants encounter ties or instances where they find both stimuli equally preferable, they are instructed to choose between them randomly. This method is mainly preferred because of its relative simplicity compared to other methods and because it is better at finding differences between images. The task typically indicates the preferred option from each pair of stimuli rather than assigning a quality score to each stimulus.

The results can be presented as a winning frequency matrix that illustrates the relative frequencies with which each stimulus is preferred over the others. The Binomial Sign test can be computed using the frequency matrix to show the statistical significance of the result obtained. To account for the possibility of type I errors resulting from multiple conditions, a Bonferroni correction (Bland et al. 1995) is applied. Additionally, the Z-score (Engeldrum 2001) is computed based on Thurstone's law of comparative judgment (Thurstone 2017). It is a statistical measure that indicates how far a particular observation is from the mean in terms of standard deviation and allows one to draw statistical inferences about the differences between the two items being compared. To compute the z-score, the frequency matrix obtained is transformed into a percentage matrix by dividing the frequency values by the total number of observations. Subsequently, a Logistic Function Matrix (LFM) is generated using an Equation (2.2) (Engeldrum 2001).

$$LFM = \ln \frac{(f + c)}{(N - f + c)} \tag{2.2}$$

Where, $f$ represents the value from the frequency matrix, $N$ represents the total number of observations, and $c$ is an arbitrary additive constant. Typically, the value of $c$ is set to 0.5, which is the most commonly used value.

The computation of LFM is a method used to model the decision-making process in a pair comparison experiment. By analyzing the observed frequencies of preference, LFM helps identify the relative strengths of preference for each stimulus and their

influence on decision-making. Additionally, the LFM can predict the probability that a given stimulus will be preferred over another in future trials. The LFM can be converted to z-scores by applying a scaling coefficient. This coefficient is determined by analyzing the relationship between the inverse of the standard normal cumulative distribution for the percentage matrix and the values in the LFM. This analysis is carried out through linear regression and the scaling coefficient is obtained from the slope of the regression line. Finally, a z-score matrix is obtained by multiplying the LFM matrix with the scaling coefficient, and it is shown with a 95% confidence interval.

There are standards (BT.500-14 2020; P.910 2022) that provide guidelines and methodologies for conducting and evaluating subjective quality assessment. However, the implementation of subjective assessments can incur high costs, require considerable time and effort, and pose challenges when attempting to scale them up for large datasets. Objective quality assessment methods, commonly formulated or trained using data from subjective assessments, can help mitigate some of the limitations of subjective assessment by automatically estimating visual quality using mathematical models that can predict the quality evaluation of an average human observer. There are three primary approaches to achieving an objective image quality assessment. The first approach involves using full reference metrics, where a reference image (an undistorted version of the same image) is compared with the distorted image to calculate quality scores. These metrics consider various factors such as structural similarity, pixel-level differences, or perceptual features. The detailed survey on the full reference-based metrics can be found in (Pedersen and Hardeberg 2012). The second approach is based on reduced reference metrics, which utilize a subset of information from the reference image to estimate the quality of the distorted image. This approach reduces computational complexity while maintaining reasonable accuracy. Some of the reduced reference base metrics are found in (Gao et al. 2009; Tao et al. 2009). The third approach is called no-reference metrics, where the quality assessment is conducted without any reference to the original image. In this approach, the quality of an image is assessed solely on its characteristics, such as local features, global features, perceptual features, or statistical properties. For a comprehensive understanding of the methods developed for no-reference-based image quality assessment, please refer to (Kamble et al. 2015). Each approach has its strengths and limitations, and the choice depends on the specific requirements of the application and the available resources.

### 2.4.1 Quality Assessment in the Cultural Heritage Context

Digitalization of CH requires substantial investment in time, effort, and resources. By adhering to minimum quality standards, the resulting digital surrogates gain enhanced long-term value, ensuring their enduring significance and practicality for future generations. Within the CH field, limited standardized tools and guidelines are available to analyze image quality. These tools employ various test charts (also referred to as targets) and analysis algorithms designed to measure image quality

characteristics such as resolution, sharpness, color accuracy, noise, distortion, and artifacts, contributing to the establishment of sustainable digitization practices that facilitate the long-term preservation of digital materials.

Numerous CH organizations and vendors in the CH community worldwide have widely accepted three global standards, namely the Federal Agencies Digitization Guidelines Initiative (FADGI) (FADGI 2010), Metamorfoze (Dormolen 2012), and ISO 19264-1 (ISO19264 2021), as the key directives for incorporating objective image quality assessment into the process of digitizing two-dimensional CH artifacts. FADGI is a collaborative effort among various federal agencies to establish guidelines and best practices for digitizing CH materials. It introduces a categorization framework that comprises four levels that are used to assess the quality of the imaging. These levels are represented by star ratings, ranging from 1 to 4. A higher number of stars signifies an increased level of consistency in the captured images. Metamorfoze is a European preservation guideline developed by the National Library of the Netherlands specifically for safeguarding paper-based heritage. The Metamorfoze system establishes three quality levels: Metamorfoze, Metamorfoze Light, and Metamorfoze Extra Light, with Metamorfoze representing the highest level of quality. The main distinctions between these levels revolve primarily around the acceptable tolerance for color accuracy.

Although the FADGI and Metamorfoze guidelines share conceptual similarities and use similar image quality attributes, they differ in terms of test charts used, tolerance, and evaluation algorithms. Consequently, direct interchangeability between these guidelines is not possible. To address this problem and mitigate possible confusion, extensive efforts were made to harmonize the two guidelines, culminating in the development of the ISO standard ISO 19264-1:2021. This standard establishes three distinct quality levels (A, B, and C) for imaging, with Level A denoting higher image quality. These three levels align conceptually with both the FADGI star system (4, 3, and 2 stars) and the Metamorfoze three-tier system. Comparison of these guidelines in terms of their quality ratings is presented in Table 2.1.

**Table 2.1:** Comparison of FADGI, ISO 19264-1 and Metamorfoze quality rating systems for digitizing CH artificats (imageaccess 2023).

| Description | FADGI | Metamorfoze | ISO 19264-1 |
|---|---|---|---|
| High Quality | 4 stars | Metamorfoze | Level A |
| Medium Quality | 3 stars | Metamorfoze light | Level B |
| Low Quality | 2 stars | Metamorfoze extra light | Level C |
| Very Low Quality | 1 stars | No equivalent | No equivalent |

In the field of HSI, instrument manufacturers often employ different calibration methods and provide specifications in varying formats. Moreover, discrepancies arise in the definitions of terms used in different hyperspectral instrument manufactures.

As a result, there is a pressing need for a unified standard that can establish common guidelines and definitions. To date, there are no published standards for assessing spectral imaging. However, certain organizations have initiated efforts, focusing particularly on establishing standards for HSI (Eckstein et al. 2021). The National Institute of Standards and Technology (NIST) is currently developing a standard titled 'Hyperspectral Imaging Standards' (NIST 2015) that focuses on various aspects of HSI. This ongoing standard aims to encompass performance specifications, calibration standards, data formats, terminology, and best practices for HSI. Likewise, the Institute of Electrical and Electronics Engineers (IEEE) Geoscience and Remote Sensing Society (GRSS) launched an initiative in 2018 known as 'Project 4001 (P4001)' (P4001 2018). This endeavor involves the establishment of a hyperspectral Working Group under the IEEE's Standards Association. The primary focus of this working group is to address key aspects such as terminology, data structures, and characterization and testing methods within the field of HSI.

# Chapter 3

# Summary of the Articles

This chapter summarizes the articles incorporated in this thesis, comprising five research articles. Each article is accompanied by a concise summary of its motivation, methodology, and results, highlighting the main aspects of the conducted research. For more detailed information, refer to the respective manuscripts, which can be found in the second part of this thesis.

## 3.1 Article A: Influence of Acquisition Parameters on Pigment Classification using Hyperspectral Imaging

D. J. Mandal, S. George, M. Pedersen, and C. Boust (2021). "Influence of Acquisition Parameters on Pigment Classification using Hyperspectral Imaging." In: *Journal of Imaging Science and Technology* 65.5, 050406-1–050406–13

### 3.1.1 Objective

In ch, the classification of pigments in artworks, such as paintings, is essential. This is of significant importance to curators and conservators, as it enables them to gain deeper insights into an object and to know its historical values (Mayer 1966; Stuart 2007). For several years, hyperspectral imaging technology has been employed for pigment classification, offering significant potential in its scientific analysis (Bai et al. 2017; George et al. 2019). However, the acquisition of hyperspectral images presents several challenges. Various parameters, including focus, signal-to-noise ratio, and illumination geometry, can impact the quality of acquired hyperspectral data (Foster et al. 2019; Pillay, Hardeberg, et al. 2019). To better understand the impact of these acquisition parameters on pigment classification application, we investigated their effects in a mockup using HSI in the VNIR region in this study.

### 3.1.2 Methodology

We collected hyperspectral data in a region of 400 to 1000 nm with 186 spectral bands, having a spectral sampling of 3.26 nm. The VNIR camera was used with a 30 cm cylindrical lens to capture 1800 spatial pixels across a line, which covers an approximate field of view of 86 mm. The pigment mockup was placed on a translation stage setup (Figure 3.1), with optical fiber-based illumination having a geometry of 45°- 0°- 45°, where 0° indicated the camera angle with normal. We adjusted the camera-to-mockup distance in 2 cm increments from the reference focus point to analyze the effect of varying focus. Due to the arrangement of the setup, it was convenient to move the camera in an upward direction, as shown in Figure 3.1. For the signal-to-noise ratio (SNR), each line in the mockup was scanned 'N' times (ranging from 1 to 8) and averaged before proceeding to the following line. Data were acquired with varying integration times by keeping the saturation of the acquired pixels within the limit of 10 - 85%. We also analyzed data acquired at illuminant angles of 30 and 60 degrees. A supervised-based approach using the Spectral Angle Mapper (SAM) algorithm was applied for classification because it is one of the most commonly applied classification algorithms for spectral matching.

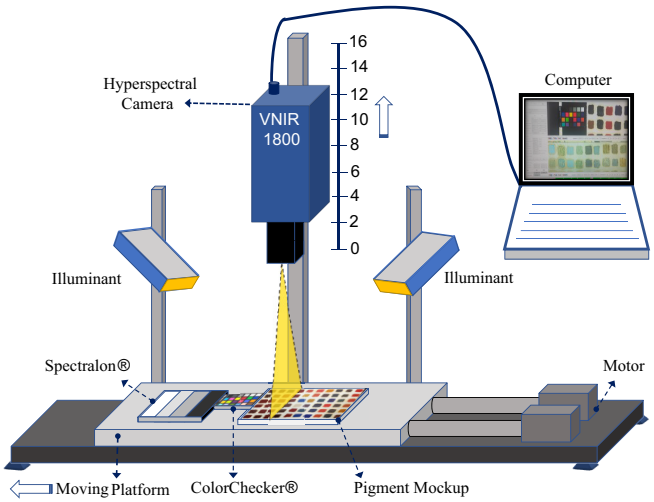



**Figure 3.1:** Layout of the HSI system used in the experiment. The object was illuminated using a 150 W halogen-based SmartLite 3900e by Illumination Technologies, Inc., with light guided through an optical fiber. The camera was moved upward from 0 to 16 cm with a step size of 2 cm; 0 indicates the optimal focus distance point (i.e., 22 cm from an object to the lens).

### 3.1.3 Results

The study found that when the mockup is placed away from the optimal focus point, the resulting image becomes blurry, functioning as a low-pass filter, and consequently increasing classification accuracy. However, beyond a certain distance, the classification accuracy starts to decrease. The influence of SNR and integration time on classification accuracy was relatively minimal compared to the impact of focus distance, likely due to lower noise levels in controlled laboratory environments. The mockup used in the experiment had an uneven surface of pigment patches, resulting in significant variations in the acquired spectrum within the same patch. Furthermore, alterations in the illumination angle also affected the accuracy of the classification. Implementing an equalization filter effectively mitigated the noise in the obtained spectrum, particularly at both ends of the wavelength range within the VNIR region.

## 3.2 Article B: An Experiment–based Comparative Analysis of Pigment Classification Algorithms using Hyperspectral Imaging

D. J. Mandal, M. Pedersen, S. George, H. Deborah, and C. Boust (2023). "An Experiment–based Comparative Analysis of Pigment Classification Algorithms using Hyperspectral Imaging." In: *Journal of Imaging Science and Technology* 67.3, pp. 030403-1–030403-18. DOI: 10.2352/J.ImagingSci.Technol.2023.67.3.030403

### 3.2.1 Objective

In the existing literature, numerous supervised algorithms for HSI have been developed, primarily for remote sensing purposes, such as mineral identification (Melgani et al. 2004). However, only a limited number of these algorithms have been implemented in the CH domain, specifically for pigment classification tasks in artworks such as paintings (Mandal, George, et al. 2021). The shape of each pigment spectrum is distinctive because every material has a different chemical composition and an inherent physical structure (Shaw et al. 2002). These distinct features make spectral matching approaches widely utilized for pigment classification, where the similarity between two spectra at a specific pixel in an image is evaluated. The pigments can be accurately classified on the closest match by comparing the spectra with a set of reference spectra. In the CH context, HSI acquisition is commonly carried out under controlled laboratory conditions, where the camera is positioned relatively close to the object and the illumination is carefully regulated. On the other hand, remote sensing data collection occurs under natural illumination conditions, from a greater distance, which introduces atmospheric effects and temporal variations in illumination. As a result, the inherent differences between remote sensing and

CH applications mean that classification algorithms specifically designed for remote sensing may not be equally applicable or well-suited for CH purposes.

Machine Learing (ML)-based classification methods have been popular and widely used in many applications in recent years. Support Vector Machine (SVM) and Convolutional Neural Networks (CNN) are a few of the ML approaches used for classification tasks that learn spectral features more effectively using deeper layers and can give us higher classification accuracy than traditional algorithms (Melit Devassy et al. 2020; Pouyet et al. 2021). The difference between various classification algorithms lies in their ability to account for differences in the shapes and magnitudes of the two spectra being compared. In this study, we evaluated the performance of eight supervised-based algorithms, namely SAM, Spectral Correlation Mapper (SCM), Spectral Information Divergence (SID), Spectral Similarity Scale (SSS), and the hybrid combinations of SID–SAM and SID–SCM. We also used the Jeffries–Matusita (JM) distance function combined with SAM (JM-SAM). Likewise, three ML models, SVM, Fully Connected Neural Network (FC-NN), and 1Dimentional-CNN (1D-CNN) for pigment classification of a mockup using HSI.

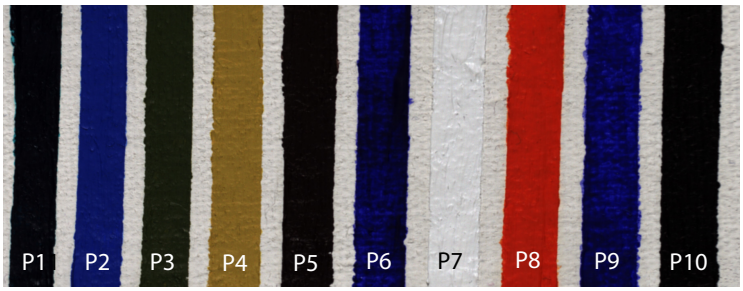



**Figure 3.2:** A flat section of a pigment mockup illustrating various pigments: P1:Veridian, P2:Cerulean blue, P3:Green earth, P4:Yellow ochre light, P5:Burnt umber, P6:Ultramarine blue deep, P7:Lead white hue, P8:Genuine vermilion, P9:Cobalt blue deep and P10:Ivory black. Colors are approximated as RGB rendering using spectral python for bands 75, 46, and 19 of the datacube.

### 3.2.2 Methodology

As shown in Figure 3.2, a pigment mockup was prepared and used in a laboratory environment; device specification, acquisition setup, and data preprocessing steps were similar to the study mentioned in Section 3.1 (Article A). The selection of pigments for this mockup was based on extensive research articles in the field of CH, spectrum characteristics, and expert consultation. Selecting the appropriate threshold value for classification algorithms is critical as it may vary depending upon the application (H. Li et al. 2014). We used an empirical approach similar to (Carvalho Júnior et al. 2011) for computing the optimal threshold for each of these supervised algorithms and evaluated their accuracy using a confusion matrix.

### 3.2.3 Results

The experimental findings indicate that the ML algorithms performed better than the supervised-based algorithms utilized in this study. Supervised-based algorithms exhibit limitations in scenarios where pigments have nearly identical spectra (e.g., P6 and P9 in Figure 3.3) or when the magnitude of the spectrum is very low (e.g., in Figure 3.3, P10 with reflectance factor below 0.05). In situations involving nearly identical spectra, the SCM proves to be a more suitable measure than the SAM; additionally, the classification accuracy of certain pigments is influenced by the threshold value chosen. Algorithms based on spectral distance, such as ED, SSS, and JMSAM, had the lowest classification accuracy.

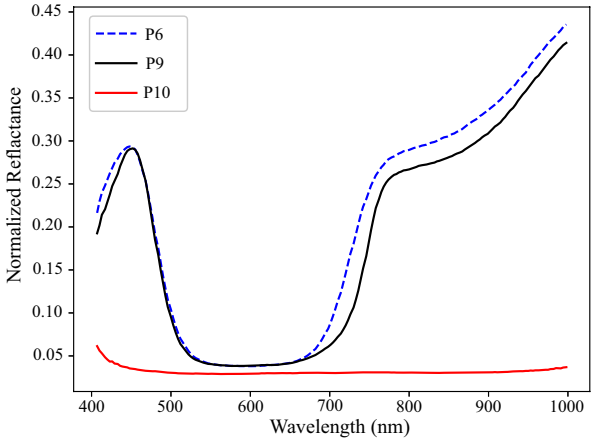



**Figure 3.3:** Normalized reflectance spectrum for pigments; P6:Ultramarine blue deep, P9:Cobalt blue deep and P10:Ivory black.

Although more effective, ML-based algorithms require substantial amounts of training data and hyperparameter tuning, resulting in increased computational costs and complexity. On the other hand, supervised-based algorithms offer simplicity and ease of computation. Therefore, supervised-based algorithms, such as SCM and SAM, might be a good fit for the classification task for pigments with less complex spectra.

## 3.3 Article C: A Comparison of Pigment Classification Algorithms on Non–Flat Surfaces using Hyperspectral Imaging

D. J. Mandal, M. Pedersen, S. George, and C. Boust (2023). “Comparison of Pigment Classification Algorithms on Non–Flat Surfaces using Hyperspectral Imaging.” In: Manuscript under revision in Journal

### 3.3.1 Objective

Artwork such as paintings is not confined to two-dimensional canvases or boards, which means that they are not always flat. The addition of relief, which introduces 2.5D or 3D to the artwork, is also an important consideration (Rubens n.d.; Townsend 1995). Several factors can contribute to this third dimension (Barrett et al. 1995; Baxter et al. 2004); for example, the morphological textures of brushstrokes on the painted surface (Fu et al. 2018), a thick layer of pigments applied by many renowned artists to their artwork to create depth (impasto technique) (Elkhuizen et al. 2019; Gonzalez et al. 2019; Plisson et al. 2014). The geometry of a relief raised from a brush painting, impasto techniques, or any other factors may affect how light interacts with the surface, affecting the spectral signature captured by the hyperspectral sensor for a given pixel. To our knowledge, no previous research has been conducted on pigment classification using HSI concerning an elevated surface. Therefore, in this study, we compare different spectral classification techniques that employ deterministic and stochastic methods, their hybrid combinations, and machine learning models for an elevated mockup to determine whether such topographical variation affects classification accuracy. As demonstrated in Section 3.2 (Article B), ML models have improved accuracy for pigment classification. However, these models require large datasets to train effectively. This poses a significant challenge in the CH domain because the data for ML applications are often inadequate. This challenge can be addressed by generating synthetic data using data augmentation. In this regard, the impact of data augmentation techniques on the effectiveness of ML models for CH applications was also analyzed in this study.

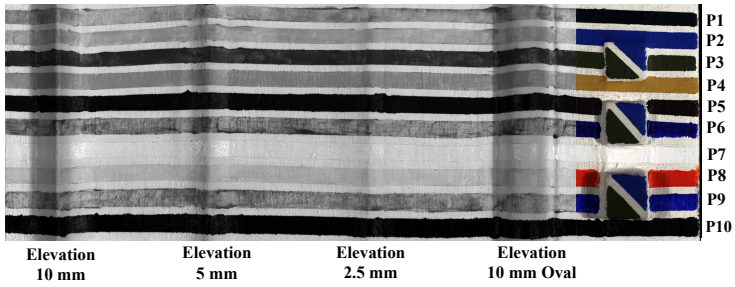



**Figure 3.4:** Pigment mockup used in the study, consists of ten pigments labeled P1 to P10. The mockup includes a flat region and three different elevation levels (2.5mm, 5mm, and 10mm). For better visualization of the elevation, the left part of the image is presented in grayscale, captured at 998 nm in the NIR region, while the right side displays a color image generated using bands at 640, 551, and 458 nm.

### 3.3.2 Methodology

In the laboratory, a pigment mockup, as illustrated in Figure 3.4, was created using a 3D printed base; It consisted of different elevated levels, including a flat surface and

regions raised to levels of 2.5mm, 5mm, and 10mm from the base. Similar to our previous study (Section 3.1 and Section 3.2), hyperspectral data were acquired in a laboratory using HySpex VNIR-1800 and a translation stage setup. In this setup, having different elevated levels within the mockup, we aligned the height of the Spectralon® (Spectralonl 2023) tiles with the flat surface level of the mockup. We used and assessed the identical sets of algorithms used in our previous study Section 3.2 (Article B) for the classification of pigments. For data augmentation, we introduce four attributes to the spectrum: offset, multiplication, Gaussian noise (Davenport et al. 1958), and speckle noise (Boyat et al. 2015). An example of the implementation of these attributes is shown in Figure 3.5. The offset was varied ±(0.0001 to 0.1 with a step size of 0.001) times the standard deviation of the training set. The multiplication was done with 1±(0.0001 to 0.1 with a step size of 0.001) times the standard deviation of the training set, and the two different noises, Gaussian distributed additive noise and speckle, which is a multiplicative noise, were added ten times with a variation of 0.00001 and 0.000001 respectively.

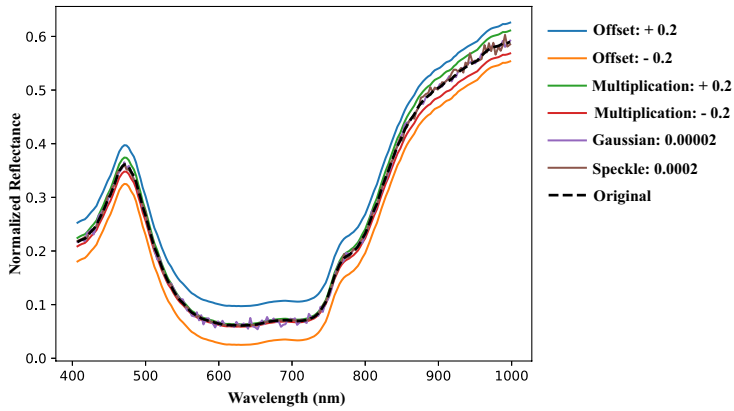



**Figure 3.5:** Data augmentation of a spectral signal; produced with an addition of offset and multiplication to data by a factor of 0.2; Gaussian noise and speckle noise are added to the data with variations of 0.00002 and 0.0002, respectively.

### 3.3.3 Results

From our experimentation, we observed that elevation itself does not significantly affect the algorithm for classification accuracy, but the formation of shadows caused by elevation can have a significant impact. ML models 1D-CNN and SVM perform better than the eight supervised algorithms, the least accurate being the ED algorithm. Among the eight supervised algorithms, SAM and SCM performed better, while the algorithms employing hybrid approaches did not perform well overall. Although there are variations in accuracy across different algorithms, we can discern a pattern in the classification results: the classification accuracy decreases as the elevation increases.

We also observed that irrespective of the algorithms used, classification accuracy also depends on the type of pigments; for example, regarding misclassification, some pigments are affected more than others. In addition, we also noticed that the choice of reference spectrum used for building the spectral library and training ML models plays a crucial role in classification accuracy. Using average reference spectra, classification accuracy was improved for almost all flat and elevated regions compared to the accuracy obtained when individual reference spectra were obtained from four different elevation surfaces.

Additionally, the results demonstrated that data augmentation played a crucial role in improving classification accuracy, particularly when multiple spectra were augmented or when an average spectrum was employed for augmentation. Figure 3.6 illustrates the classified images obtained using an SVM model under various data augmentation conditions.

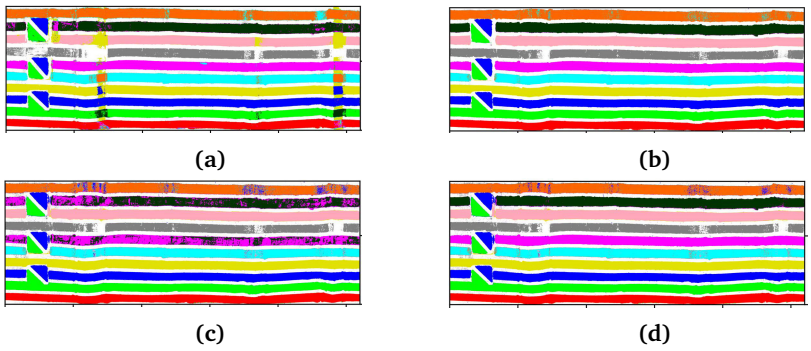



**Figure 3.6:** Classification results for SVM using different data augmentation conditions: (a) Without data augmentation, (b) With data augmentation, (c) With data augmentation using a single spectrum from the flat region, and (d) With data augmentation using a single spectrum from the averaged region.

## 3.4 Article D: Evaluation of Text Legibility in Alternative Imaging Approaches to Microfiche Digitization

H. Deborah and D. J. Mandal (2021). "Evaluation of Text Legibility in Alternative Imaging Approaches to Microfiche Digitization." In: *Proc. IS&T Archiving*, pp. 96–101. DOI: 10.2352/issn.2168-3204.2021.1.0.22

### 3.4.1 Objective

Before the advancement of digital technologies, microforms, such as microfilms and microfiche, were the only method available to archive and conserve extensive documents. The CH sector quickly adopted this technique to capture its collections for preservation, accessibility, and distribution. Today, many historical collections are only available in microfiche (Garcia-Spitz 2017), making their preservation an important task. Microfiche cannot be read directly by the human eye and requires a special device for enlarging, printing, and scanning the microforms into readable formats. Microfiche readers may not be as available as before, since digital technologies have mainly replaced them; additionally, such readers are made for reading and not for digitization or data collection. As digital technologies continue to be adopted, digitizing microfiche presents numerous benefits (Hirtle 2002). Digitization makes it easier to store and manage microfiche collections, and creating digital copies can also protect against loss or damage. However, it is crucial to ensure the quality of the digitized images to meet user objectives. In this study, we evaluated the performance of two imaging devices as alternatives to traditional microfiche readers.

### 3.4.2 Methodology

We employed a microform reader, a flatbed scanner, and an in-house film scanner equipped with a monochrome camera and a macro lens to scan the microfiche. The resulting digital images were post-processed using a median filter. The microfiche used in this study mainly comprises textual content; consequently, the digitized version of this content should be easily readable. Thus, we considered legibility as the evaluation criterion and optical character recognition (OCR) to read the text. To evaluate quality, we used the Levenshtein edit distance (Levenshtein 1966) metric to measure the text similarity between the obtained text and the ground truth. In Figure 3.7, a flow diagram is presented, illustrating the assessment of text legibility in the context of microfiche digitization.

### 3.4.3 Results

The study found that an alternative imaging device performs better than traditional microform readers in terms of text legibility. Additionally, the use of a median filter for noise removal and smoothing improved legibility. Among the imaging devices tested, the flatbed scanner with 4800 dpi was the most suitable for providing the highest quality computer legible texts. We observed that the images acquired from the microform reader exhibited a noisy and granular background, whereas those from the flatbed scanner displayed a smoother background. However, the flatbed scanner images suffer from lower contrast and are less sharp. We applied a median filter to the microform reader images to address this. The results showed a notable improvement in legibility after applying the filter.

We comprehensively compared by carefully evaluating the suitability of applying

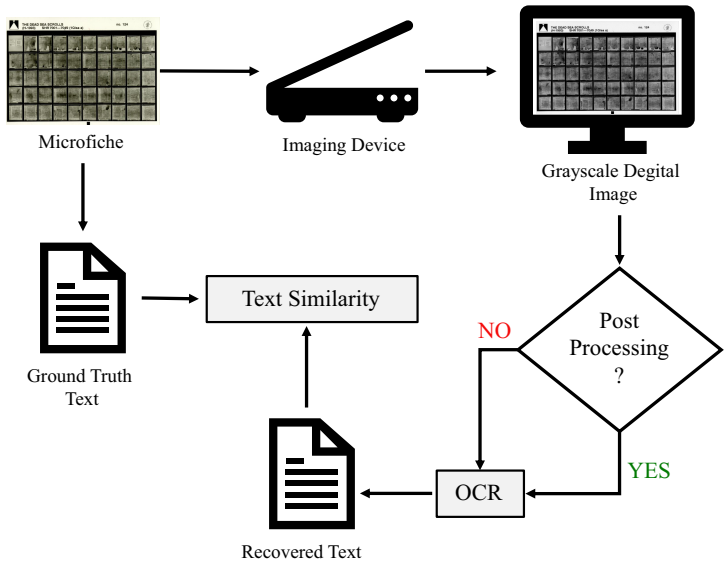



**Figure 3.7:** Experimental flow diagram illustrating the evaluation of text legibility using optical character recognition (OCR); three imaging devices are compared: a microform reader, a flatbed scanner, and an in-house film scanner.

median filters with different kernel sizes to images acquired from three imaging devices. Our findings indicated that the flatbed scanner operating at 4800 dpi still provided the best overall performance. However, by applying a median filter on the microform reader images, the result was significantly improved, bringing them closer to the quality of the flatbed scanner images captured at 4800 dpi. The result showing the cumulative Levenshtein edit distance for images obtained from three different imaging devices, before and after applying the median filter, is shown in Figure 3.8.

## 3.5 Article E: Subjective Quality Evaluation of Alternative Imaging Techniques for Microfiche Digitization

D. J. Mandal, H. Deborah, and M. Pedersen (2023). "Subjective Quality Evaluation of Alternative Imaging Techniques for Microfiche Digitization." In: *Journal of Cultural Heritage*. Accepted for publication

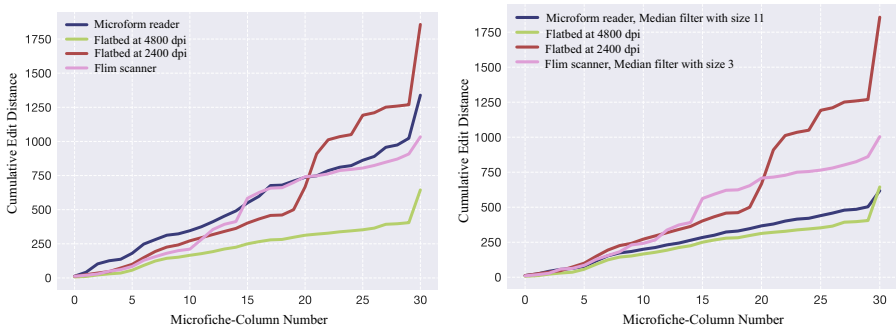



**Figure 3.8:** Comparison of Levenshtein edit distance for images acquired using three imaging devices, shown in a cumulative manner along the x-axis; (a) without median filters: The result shows that the flatbed scanner at 4800 dpi has the best performance, and (b) Both microform reader and film scanner images are combined with median filters as post-processing steps.

### 3.5.1 Objective

Microfiche materials extend beyond written texts, as they encompass a variety of records, not limited solely to textual documents. There is a need to explore the quality assessment of digitized microfiche containing various types of content, such as photographs and handwritten text. For microfiche digitization, subjective experiments can help to assess the perceived quality of the digitized images by human observers. This is important because humans are the ultimate consumers of digitalized material. Image enhancement is a frequently used method to improve the visual quality of images by adjusting their attributes (Singh et al. 2014). Among the quality attributes of an image, contrast is considered crucial (Pedersen, Bonnier, et al. 2010), and its enhancement is believed to improve the visual quality of most natural images (Gu et al. 2013). In this study we incorporated three different microfiche materials containing typewritten text, photographs of natural scenes, and photographs of ancient handwritten fragments. We applied contrast stretching as a postprocessing method to evaluate its effectiveness in improving the quality of microfiche materials using subjective quality assessment. Our results showed that the reproduction of alternative devices was preferred over that of a traditional microfiche reader. Furthermore, our results demonstrate that image enhancement techniques significantly improved image quality. This study suggests that alternative imaging devices may be a viable option for digitizing microfiche and improving access to historical collections.

### 3.5.2 Methodology

In this study, we used microfiche materials from three distinct sources. These sources include microfiches provided by a handbook for evaluating microfiche readers (Hawken

1975a), the Allegro Qumran Collection on Microfiche (Brooke et al. 1999), and the Dead Sea Scrolls on Microfiche (Tov et al. 1995). Figure 3.9 shows an example of the image obtained from each microfiche. The imaging device and preprocessing steps used in the experiment were identical to the one mentioned in Section 3.4.2, and we designed a subjective experiment as a force-choice pair comparison (David 1988), with no tie option, using a web-based tool called QuickEval (Ngo et al. 2015) and in a controlled environment. There were a total of 21 observers, including one expert. The expert observer had experience with imaging, old/historical manuscripts, and microfiche, all of which were used in the experiment. The results were analyzed using two statistical methods: z-score analysis and the sign test. These analyses were conducted using an open-source Python platform.


**Figure 3.9:** Examples of microfiche material images used in the experiment acquired through a in-house film scanner. (a) Text, (b) Natural Scene, and (c) Fragments.

### 3.5.3 Results

The results showed that alternative imaging devices were preferred over traditional microform readers, indicating that they could be a viable option to digitize microfiche and improve access to historical collections. The result of applying image enhancement techniques also significantly improved the image quality for all three categories of microfiche. We performed a separate analysis distinguishing between the expert and the naive users. The results revealed that the naive observers' preferences strongly aligned with the expert's. Furthermore, we analyzed the general preferences of observers regarding enhanced and non-enhanced images across different devices, regardless of the microfiche type. As illustrated in Figure 3.10, the results indicate that most observers consistently preferred enhanced images over non-enhanced images across all imaging devices.

In addition, a notable observation from the results found was that the observers' ability to differentiate between devices is becoming less distinct as we compare microfiche with text and natural scene; this distinction is even less for microfiche with fragments. A notable observation from the results was that the ability to differentiate between devices is becoming less distinct as we compare microfiche with text and natural

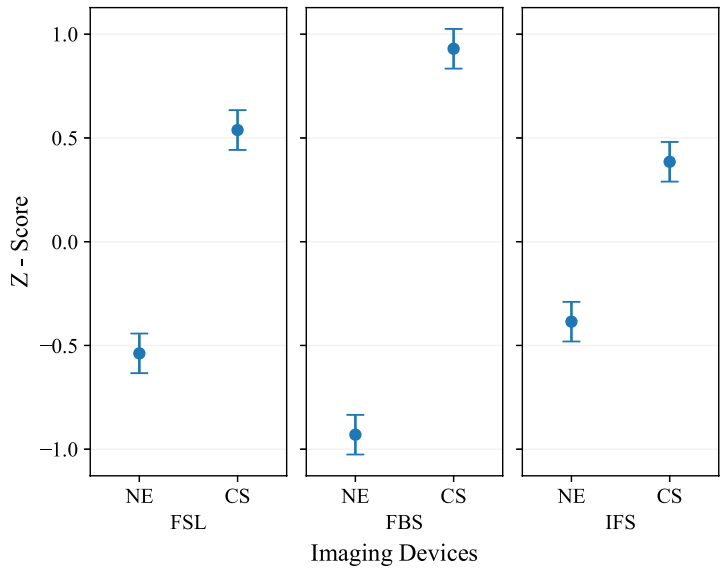



**Figure 3.10:** Mean z-scores with 95% confidence intervals for non-enhanced (NE) versus contrast-stretched (CE) images across three imaging devices; microform reader (FSL), flatbed scanner (FBS), and in-house film scanning (IFS); higher z-score value is better.

scene; this distinction is even less for microfiche with fragments. In future research, it would be beneficial to conduct objective image quality assessments to evaluate different microfiche quality attributes, such as contrast and sharpness. This could involve exploring existing image quality metrics to determine their relevance for time-efficient evaluations of microfiche image quality.

# Chapter 4

# Discussion

In this chapter, we look from a broad perspective at the research presented in this thesis, examining its impact and contributions. It is divided into two parts: The first part addresses the significance of the research findings of each article summarized in Chapter 3, their alignment with the research objectives described in Section 1.2, and their limitations. The second part highlights the specific contributions made by this research to the realm of CH imaging and discusses its potential applications in other domains, specifically in the context of quality assessment.

## 4.1 Quality Assessment for Spectral Imaging

Article A, B and C contribute to the quality assessment of digitizing CH artifacts, such as paintings using HSI in the Visible and Near-Infrared (VNIR) spectrum. Spectral quality here is defined in terms of the accuracy with which the system can identify pigments for the classification task. Different spectral metrics are used to measure the quality, which computes the similarity between the obtained spectrum and its corresponding ground-truth spectrum, relying on empirically computed threshold values. These articles specifically address the first three research questions, RQ1, RQ2 and RQ3, which are discussed below.

### 4.1.1 RQ1: How do acquisition parameters impact the accuracy of pigment classification in digitizing artworks using HSI ?

Despite significant utilization of HSI in CH, there are still important challenges in delivering high-quality spectral data related to acquisition. In CH, the acquisition is usually carried out in a close range, and the quality of the data obtained can be influenced by several acquisition parameters of the HSI system. Article A, provides analysis on how some of these parameters can influence the quality of the hyper-spectral data obtained, particularly for the task of pigment classification using a pigment mockup. For artwork such as paintings, pigments used have heterogeneous

structures; in addition, many artworks have been varnished, which can significantly affect gloss level; thus, obtained data can be influenced due to illumination geometry. The surfaces of paintings are often uneven, and acquiring data within a focus can be difficult. CH objects are susceptible to temperature, artifacts exposed under illumination for a long time can cause damage by changing their material property, and when the integration time is increased, the sensor collects more of both the signal and the noise, resulting in a higher overall noise level. Therefore, in Article A, we selected these four key acquisition parameters: focus distance, signal-to-noise ratio (SNR), integration time, and illumination geometry. The results were analyzed in terms of classification accuracy using the Spectral Angle Mapper (SAM) as a metric.

When the focus distance is changed, which involves moving the camera away from an object, it leads to variations in the magnitude of the spectrum, while the shape of the spectrum remains relatively constant. The variations in magnitude are particularly noticeable in the wavelength range of 600-1000nm for most of the pigments used. Furthermore, for a given acquisition setup, the spectrum of a pigment varies significantly across different spatial locations within the same pigment patch. This variability is influenced by both the mixed pigment concentration and the non-uniformity of the pigment patch. Some pigments show higher variability in their spectra compared to others. As a result, certain patches are classified more accurately than others. The accuracy of the pigment classification initially improves when the camera is moved away from an object. However, as the camera continues to move further, the accuracy starts to decline. When the camera is initially moved away by a small distance, the pixels become blurred, and adjacent pixels are more likely to be averaged. This leads to an increase in the classification accuracy. However, as the camera moves further away, it goes out of focus due to its limited depth of field. Consequently, photons from the pigment patch area no longer hit the same pixels and start to hit other pixels far around its neighborhood. As a result, the classification accuracy decreases.

The increase in SNR had minimal impact on the spectrum, and consequently, the classification accuracy remained relatively constant. Typically, the spectrum obtained from a Complementary Metal-Oxide Semiconductor (CMOS) detector exhibits noise at the lower and upper wavelengths because of the lower quantum energy sensitivity of the detector at these ends. Increasing the SNR value helped to reduce this noise. One might question the purpose of increasing the SNR if it does not significantly affect accuracy. However, it is important to consider the experimental conditions under which this study was conducted. The experiment was carried out at a distance of 30 cm between the device and the object. It is reasonable to assume that if this distance is further increased, for instance, by using a 1 meter lens or longer to scan a larger painting, the observed variation in the spectrum may play a more significant role. Therefore, increasing the SNR in such scenarios could prove beneficial in reducing noise and maintaining accuracy.

In the given HSI device setup, increasing the SNR requires scanning the same line of pixels multiple times and averaging the spectrum obtained. However, this approach has a drawback as it prolongs the object's exposure to illumination, which can be destructive to light-sensitive artifacts. To address this problem, a potential solution is to use an equalizer filter. Through our experiment, we observed that the equalizer filter effectively suppresses noise at both ends of the spectrum. This enables improved noise reduction without the need for multiple scans or prolonged exposure times, thus minimizing the potential damage to light-sensitive artifacts. A similar observation was made when the integration time in terms of its impact on the spectrum variation and classification accuracy. However, the same trend did not hold for changes in the illumination angle. The highest classification accuracy was achieved at the standard 45-degree angle, while the accuracy varied with different angles. The least variation in accuracy was observed at a 30-degree angle, while a more significant variation was observed at a 60-degree angle. Additionally, it was noted that the non-homogeneity in the surface of pigment patches influenced this variation.

### Limitations

Although Article A offers an extensive examination of the acquisition parameters for HSI in pigment classification, it limits itself to the pigment mockup used in this study. Pigment patches are separated from each other, which does not reflect the typical arrangement of elements in real paintings, where they are usually in close proximity. Consequently, when there is a variation in focus distance, adjacent pixels containing different pigments may be erroneously classified due to changes in spatial resolution. A further detailed study is needed to analyze the nonhomogeneity of pigment surfaces resulting from factors such as brushstrokes, variations in layer thickness, and pigment compositions.

In addition, it is important to acknowledge that using an unvarnished mockup in the study may restrict the applicability of findings to real-world scenarios, where paintings are usually varnished, and specular reflection may arise due to varying illumination geometries. The classification algorithm used in the study is one of the most widely applied algorithms for spectral matching. However, it is crucial to explore additional supervised algorithms as they may offer improved performance in the previously discussed scenarios. Evaluating the performance of different algorithms across various pigment types can contribute to a more robust and effective approach to classifying pigments.

## 4.1.2 RQ2: What is the effectiveness of existing classification algorithms for pigment classification in artwork, and how can they be optimized ?

Pigment classification with HSI commonly employs supervised classification algorithms. These algorithms compare the spectrum of a region of interest with a spectral

library by applying predetermined threshold values. In remote sensing applications, numerous classification algorithms are prevalent for hyperspectral data. Some of these algorithms are also being applied directly or with slight modifications in other domains like medicine, food, agriculture, and many more. It is important to note that the effectiveness of each algorithm varies depending on the specific application domain in which it is applied. The CH domain differs from remote sensing in terms of both the acquisition setup and the interpretation of algorithmic results for end-user tasks. In CH applications, HSI acquisition is usually performed under controlled laboratory conditions with a relatively small distance between the camera and the object, allowing control over illumination and its geometry. This contrasts with remote sensing, where such controlled conditions are not feasible.

Furthermore, certain algorithms used in remote sensing are insensitive to variations in the intensity (magnitude) of spectra, which is advantageous for mitigating the effects of temporal illumination changes in the data. However, these algorithms may be less effective in CH applications, where magnitude measures are crucial. Therefore, it becomes necessary to investigate the effectiveness of these algorithms, specifically in the CH domain. Addressing this necessity, Article B compares eight supervised-based algorithms, namely Spectral Angle Mapper (SAM), Spectral Correlation Mapper (SCM), Spectral Information Divergence (SID), Spectral Similarity Scale (SSS), hybrid combinations of SID–SAM and SID-SCM, and the Jeffries–Matusita (JM) distance function combined with SAM (JM-SAM), including traditional machine learning (ML) models, Support Vector Machine (SVM), and two deep learning models, Fully Connected Neural Network (FC-NN) and 1Dimentional-CNN (1D-CNN), for pigment classification of a mockup using HSI.

The performance of different supervised algorithms varied when classifying pigments. The effectiveness of these algorithms depended on the properties of the spectra of the pigments. Specifically, most supervised-based algorithms exhibited poor performance when dealing with similar spectra that only differed slightly in magnitude. Similarly, pigment spectra with low magnitudes across the entire range of wavelengths also resulted in inadequate performance. Distance-based algorithms, namely ED, SSS and JM-SAM, showed weakest performance overall. It was observed that, particularly for a similar spectrum, the accuracy was below 50%. The SID algorithms use a divergence measure to match the similarity between the spectra, and the smaller the divergence value, the more likely the pixels are similar. It was observed that SID and its hybrid combination did not perform well for a few of the pigments, and the reason was due to the threshold value. The empirically selected threshold was sufficient to classify the eight out of ten pigments, while for the rest of two pigments, this was above the required threshold value for most of the pixels. One can change this value to get most of the pixels classified, but this will result in higher misclassification. When comparing the supervised algorithms, SCM and SAM performed better on average.

In contrast, all ML-based models performed better for pigment classification compared to supervised-based approaches. Such models often require large datasets to capture nonlinear interactions between features, making them more versatile for diverse and challenging classification problems. However, it is worth noting that ML models often require a number of parameters to be tuned (hyperparameters) and optimization (selecting suitable architectures) to learn such intricate features within the data and achieve optimal performance, which increases computational complexity and computing time. On the other hand, supervised-based algorithms do not require such a large training set and are, therefore, simple and easy to compute. Depending on the complexity of the spectrum and the available time and resource, algorithms such as SCM and SAM might still be a good fit for the classification task.

### Limitations

As demonstrated in Article B, the application of ML models for pigment classification yields impressive results. Like any ML problem, Article B involves a complex model and requires substantial training data. However, in the field of CH, acquiring such a dataset poses a challenge. Typically, the available database offers only a single spectrum for each type of pigment. Another limitation of Article B lies in the mockup used for the experiment. Although the pigments were carefully selected on the basis of their historical usage, spectrum characteristics, and expert consultation, they may not fully represent the full range of pigment characteristics. Therefore, expanding the study to include a broader range of pigments would be beneficial. The study highlighted the weak performance of supervised algorithms when used to classify pigments with similar spectra. An analysis of the classification results could have been included to make the research more comprehensive by employing various orders of derivatives as a preprocessing step. Furthermore, in Article B, algorithms are tested and evaluated for the mockup consisting of only pure pigments; however, in realistic painting scenarios, pigments are in mixed form. Thus, adding mixed and aged pigments to the evaluation process can be beneficial.

### 4.1.3 RQ3: How does elevation affect the accuracy of pigment classification in artworks ?

CH objects like paintings may not always have a flat surface. Several factors, such as the impasto techniques employed, that is, applying a thick layer of pigments to create depth, the structure of the canvas or board used, the morphology of brush strokes, and several other factors, can contribute to adding relief to these objects, resulting in a 2.5D or 3D appearance. The geometry of such relief of a painting plays a crucial role in how light interacts with their surfaces, which, in turn, impacts the spectrum captured by hyperspectral sensors for each pixel. Since precise identification and classification of artworks are crucial, it is essential to thoroughly investigate the influence of these factors. Therefore, in Article C, we explore the impact of topographic variations on the accuracy of pigment classification. We used a similar set of classification algorithms

as used in Article B, and this includes different deterministic and stochastic methods, their hybrid combinations, and ML models. One of the limitations highlighted in Article B was the inadequate availability of training data for ML models. To address this limitation, one possible approach is to utilize data augmentation. This technique expands the dataset by generating synthetic data that closely resembles the original using various transformations while ensuring that the semantic interpretation of the data remains unchanged. Therefore, in Article C, we also analyzed the influence of data augmentation techniques on the effectiveness of ML models for artwork applications.

The mockup used in the study comprised three distinct elevation levels: 2.5mm, 5mm, and 10mm, along with a flat surface. In general, as elevation levels increased, the classification accuracy showed a decreasing trend. We found that elevation alone did not significantly impact classification accuracy. Instead, it appeared that the formation of shadows caused by elevation was more likely to be the contributing factor. The 1D-CNN deep learning model exhibited the highest performance, followed by the traditional ML model SVM. These two models performed better than the other eight supervised algorithms in terms of overall accuracy. After the ML models, SCM and SAM showed better performance compared to the other algorithms, the least accurate being the ED algorithm. It was interesting to note that SCM and SAM performed better than SVM, especially in regions with high elevation, where SVM exhibited lower accuracy, mainly due to shadow effects. The SAM algorithm, which determines the angle between two vectors regardless of their length, is less affected by changes in spectrum magnitude. Similarly, SCM eliminates negative correlation while preserving the characteristics of SAM. Hence, these two algorithms perform better in shadow areas.

Similarly to the observation in Article B, the algorithm SID and its hybrid combination with SAM and SCM did not perform well. The effectiveness of these algorithms varied between different pigments and decreased with increasing elevation. The SID algorithm relies on the probability distribution of the spectra, and any alteration in the data distribution can impact the overall entropy value. Additionally, changes in the dataset can have varying effects on the entropy value for normal and skewed distributions. Specifically, if both distributions are shifted equally, the symmetric distribution will experience less entropy change compared to the skewed distribution. This is due to the higher predictability and lower uncertainty of the symmetric distribution in contrast to the skewed distribution.

We computed statistical measures for all ten pigments, such as mean, median, standard deviation, and spectral flatness. In cases where a pigment exhibited a noticeable difference between its mean and median values, it indicated a skewed distribution. Moreover, if the standard deviation of such a distribution exceeds its mean, it implies that the data points are more widely dispersed. Consequently, datasets with a greater spread are more affected by the same amount of shift compared to datasets that are tightly clustered around the mean. This leads to a larger relative entropy between

the two datasets. Therefore, certain pigments are more susceptible to these effects than others. This also supports why we need to set different threshold values during classification for different pigments. We also observed that for pigments having nearly flat spectrum (identical reflectance values across different wavelengths), it makes it difficult for algorithms to differentiate between these classes, resulting in misclassification, which in general lowers the classification accuracy. Additionally, we noted that the selection of reference spectra greatly influences the performance of these algorithms for pigment classification. For example, suppose that the reference spectrum is obtained from a particular region with a given elevation level. In that case, it tends to exhibit minimal variation when applied to test datasets obtained from a region with a similar elevation level. This leads to a higher proportion of pixels being correctly classified. In contrast, using reference spectra from regions with different elevation levels results in a lower number of pixels to be classified correctly. Taking an average reference spectrum from different elevated regions yields better classification results.

In our study in Article C, we augmented the datasets by incorporating four attributes into the spectrum. These attributes included offset, multiplication, Gaussian noise, and speckle noise. We created three different training datasets and used the SVM model for analysis. For the first dataset (dataset 1), we used a single spectrum from the flat region of a spectral library and then generated the training data. We considered a single spectrum again for the second dataset (dataset 2), but this time it was an average of various elevated regions. Lastly, for the third training dataset (dataset 3), we individually augmented each spectrum with the associated training datasets. Compared to the other three conditions (without data augmentation, with dataset 1 and dataset 2), the classification accuracy of an augmented dataset (dataset 3) was higher. The performance was lower when using augmented datasets created from a single spectrum (dataset1). However, it still performed better than SCM (as previously noted, SCM performed better after ML models), indicating that even with a single spectrum available, better classification accuracy can be achieved using ML models than supervised-based algorithms. The results emphasize the effectiveness of data augmentation in improving classification accuracy, particularly when multiple spectra are augmented or when an average spectrum is considered for augmentation.

## Limitations

Although Article C provides valuable information for analyzing artwork in the CH domain, there are a few limitations that should be considered. The study suggests that the use of an average spectrum from different elevations improves classification accuracy. However, obtaining such a spectrum in a real scenario can be challenging. The variation in elevation within the mockup may not adequately represent all types of 2.5D or 3D paintings; employing a 3D mockup that can provide a more realistic representation of the painting would have added greater value to the results. Similarly to Article B, this study used the same mockup, introducing some shared limitations. ML models were trained and tested on only ten pigments, but expanding the scope

to include a broader range would have been beneficial.

Furthermore, the incorporation of a mockup with mixed pigments and a varnish layer would have generalized the results of the study. The results of the study emphasized the role of shadows in limiting the performance of the algorithms used for pigment classification. Applying shadow removal techniques before performing the classification task and analyzing the results would have been advantageous. Doing so could have further supported the claim that shadow removal can improve classification accuracy.

## 4.2 Quality Assessment for Conventional RGB Imaging

Article D and E contribute to quality assessment for conventional RGB imaging for digitizing microforms in the context of CH. These articles focus specifically on addressing research question RQ4.

### 4.2.1 RQ4: What are the available imaging technologies for microfiche digitization and how can they be enhanced to improve the quality of the resulting images ?

Microfiche, a type of microform, refers to thin film sheets made of plastic that are used for storing information, such as a document or image, in a much reduced format compared to the original. Before the advancement of digital technologies, microforms were the only available way to archive and preserve large documents. The CH sector quickly adopted it to capture its collections for preservation, access, and distribution. Today, many historical collections and documents can only be found in microfiche format, making them valuable cultural heritage artifacts. On the other hand, microfiche cannot be directly read by the human eye and requires a specialized device to enlarge, print, and scan the microform in a readable format. Only a limited number of specialized archives or libraries still maintain microfiche readers to enable access to their microform collections. Furthermore, these devices can encounter various issues over time, such as yellow screens, impacting their usability. Given the importance of preserving and accessing microfiche materials, exploring alternative options that can enhance the reading experience and facilitate the efficient utilization of microfiche becomes crucial.

Microfiche materials, despite their longer life span, are prone to physical degradation caused by various factors such as exposure to light, fluctuating temperature and humidity, improper handling, and many more. Adopting digital technologies to digitize microfiche can overcome some of these limitations. Digitizing microfiche, which requires magnification, is likely to introduce attributes that can affect the final quality of the reproduction, e.g., noise, distortion, and artifacts. Quality assessment should be included as an essential part of the digitization process to ensure that the digitized object maintains completeness, fidelity, and legibility compared to

the original. Microfiche materials can suffer from poor legibility and clarity due to their reduced size, potential degradation over time, or even due to scanning device limitations. Though, the use of image enhancement techniques might help to enhance visual quality and increase readability and accessibility.

Addressing these necessity, in Article D, we evaluated the performance of two imaging devices, a flatbed scanner (professional grade scanner aimed for scanning films), and an in-house film scanner (LED based multispectral film scanner), and compared them with the traditional microform reader (commercial device for digitizing all formats of microforms). We used microfiche with text for this experiment; therefore, legibility was considered an evaluation criterion. To measure the readability of the text, we employed OCR and evaluated it using the Levenshtein edit distance metric. As highlighted in Section 2.4.1, the importance of using standards and guidelines in the context of CH. The digitization workflow carried out in Article D and Article E was aligned with the high quality rating of the Federal Agencies Digitization Guidelines Initiative (FADGI). We paid particular attention to key parameters, such as target resolution, bit depth, and file format. However, since the microfiche images were grayscale, color management was not a relevant consideration in this context.

The result of Article D showed that the flatbed scanner at 4800 dpi performed better than all other devices. The in-house film scanner performed better than the microform reader, and the flatbed scanner at 2400 dpi exhibited the lowest performance among the devices tested. We observed that images from a flatbed scanner had reduced contrast and sharpness compared to images from a microform reader; the distance obtained showed inverse relation. The discrepancy in distance values was notably higher for images obtained from the microform reader than for images from the flatbed scanner and in-house film scanner. Despite having sharper text, images from the microform reader exhibited noticeable noise and granularity in the background. We postulated that background content might be a contributing factor that limits text legibility for OCR. To test this hypothesis, we reanalyzed the results in Article D by applying smoothing techniques utilizing median filters with varying kernel sizes. The results validated our hypothesis, confirming that the background content affected the text's legibility for OCR.

To further explore the results obtained, we applied median filters to the images captured by the microform reader and the in-house film scanner, comparing them with the results from the flatbed scanner. We found that a flatbed scanner at 4800 dpi was still the best choice. However, a noteworthy observation was made: the performance of the microform reader improved considerably after applying a median filter. It surpassed the in-house film scanner and approached the quality level of the flatbed scanner at 4800 dpi. Thus, results from Article D emphasize that the two alternative devices analyzed could serve as viable alternatives, and existing microform readers can yield better image quality in terms of text legibility through the application of smoothing filters.

In Article D, the result and analysis are based on microfiche with typewritten text. Therefore, the result obtained cannot be generalized for microfiche in general, as microfiche materials may also contain handwritten texts, photographs, etc. To address this limitation, in Article E, we incorporated microfiche with photographs of natural scenes and ancient handwritten fragments, including typewritten, and performed a subjective quality assessment. As influenced by Article D, how applying filters improved the result, we planned to test the attribute contrast, which is widely recognized as a significant quality attribute used for image enhancement. Thus, we applied contrast stretching as a post-processing to find if it enhances the quality of microfiche material. The flatbed scanner at 2400 dpi did not perform well, as seen in Article D; thus, for this study, we limited ourselves to the flatbed scanner at 4800 dpi (FBS), an in-house film scanner (IFS), and a microform reader (FSL). We conducted the psychovisual experiment using a force choice pair comparison method. This method was selected because of its exceptional ability to identify variations between images. In addition, it offers the advantage of relative simplicity in design and implementation. A total of 21 observers were recruited, including an expert. Due to the difficulty in finding experts with expertise with the microfiche used in the study and the logistical constraints of physical participation, we included only one expert.

In our observation, we found that images from IFS were preferred over FSL and FBS, regardless of whether they were enhanced or non-enhanced versions. In fact, the non-enhanced version of images captured by the IFS device performed better than the enhanced version from two other devices, demonstrating its choice for digitizing microfiche. However, this analysis showed a contrasting result in user preference for microfiche with text only when compared to the result obtained using objective metrics in Article D. Despite the FBS images exhibiting lower sharpness and contrast, Article D reported a high similarity score, indicating FBS as the preferred choice. However, the results in Article E were completely different, with FBS ranked as the least preferred device. These findings emphasize the need for subjective experiments or additional objective metrics based on subjective results to evaluate microfiche digitization. Furthermore, this discrepancy further emphasizes the importance of contrast and sharpness as important attributes in microfiche digitization.

Upon analyzing the data separately for the expert and the rest of the observers, we found that the results were consistent with those obtained from the combined observations. Furthermore, we also noted that enhanced versions of images were always preferred over non-enhanced versions for all three devices and for all three categories of microfiche used, highlighting the significance of using image enhancement techniques. Another interesting observation made was that, depending on the content of the microfiche, observers' ability to distinguish between the devices was moreover affected. i.e., the ability of observers to distinguish between the devices becomes less pronounced for natural scenes compared to microfiche with text. This distinction further diminishes for microfiche with fragments. In terms of decision-making time, observers took less time on average to judge microfiche with text compared to natural

scenes and fragments.

In Article E, a survey was conducted at the end of the experiment to assess the observers' preferences regarding different predefined image quality attributes that they considered for each type of microfiche. The result suggests that, when evaluating quality, legibility is an important attribute, particularly for microfiche with text. Sharpness is another key attribute that observers prioritize when evaluating microfiche with text and fragments, whereas, for microfiche with natural scenes, contrast was the most preferred attribute.

### Limitations

Some of the limitations highlighted in Article D were addressed in Article E; however, we could have analyzed the effect of other image enhancement attributes in Article D and E. Furthermore, since the subjective experiment is always time-consuming, further analyzing some of the existing objective metrics to find if there is any correlation with the subjective score obtained would have been an important consideration. In addition, by involving more experts in Article E, the experiment would have contributed to the generalizability of the results. Furthermore, to digitize microfiche with technical drawings, it is essential to ensure that the digitized versions accurately capture the intricate details and precision of the original drawings. Therefore, considering geometric quality attributes such as distortion, scale accuracy, resolution, and more would have provided a more comprehensive understanding and broader applicability to the field.

## 4.3 Limitations and Shortcomings of the Study

The results and analysis obtained in this Ph.D. thesis have made a significant contribution to the relevant field; consequently, in this section, we aim to highlight certain limitations and shortcomings of this research work from a broad perspective.

Our study used HSI in the VNIR region to assess quality in the invisible spectrum. Although this approach offered several advantages, expanding our analysis to the Short-wave infrared (SWIR) region would have been beneficial. Several studies (Cucci et al. 2016; Delaney et al. 2016; Wu et al. 2017) have demonstrated that pigments, which may not exhibit noticeable changes in the VNIR range, can reveal distinct features and material properties in higher wavelengths, which could potentially provide valuable insights into the differentiation of pigments and a comprehensive analysis for the task of pigment classification. Moreover, the impact of acquisition parameters on the SWIR region may demonstrate distinct behavioral characteristics compared to those observed in VNIR imaging. This disparity can be attributed to the distinct characteristics of SWIR devices, such as device sensor specifications, the number of spectral bands, resolutions, and other relevant factors. Although Article A focuses primarily on the influence of acquisition parameters in the VNIR range, it

would have been advantageous to broaden the study's scope to encompass SWIR. This would provide additional insights into how acquisition parameters impact the obtained data for a given task, further enhancing the overall understanding of the quality aspects in HSI across both VNIR and SWIR regions.

Similarly, when considering quality within the visible spectrum, we limited our analysis to grayscale images to accommodate the specific properties of the microfiche. However, it is worth noting that this analysis could have been extended to include other microforms, such as films and negatives, that contain color content. By incorporating colorimetric aspects, we could have expanded the scope of our study. furthermore, due to unforeseen circumstances surrounding the pandemic, we could not get access to the real painting. By not being able to validate the findings on an actual painting, the research's potential benefits and its capacity to address real-world challenges in the CH domain remained unexplored.

The study involved data processing and the application of various classification models, including supervised algorithms and different machine learning and deep learning models, using Python as an open-source platform. However, the lack of programming expertise, specifically among conservators and curators in the CH domain, could reduce the potential application of the results. To address this limitation, developing an open-source user interface software that allows users to perform these tasks with minimal programming skills, following a user-friendly manual or even without any programming background would have been advantageous. This approach would have significantly increased the accessibility and usability of the research findings, allowing a wider range of professionals in the CH domain to take advantage of the results of the thesis.

## 4.4 Contribution to Cultural Heritage Field

Acquiring high-quality hyperspectral data is crucial to obtain accurate and reproducible spectral data for analysis and documentation of artwork. Understanding how the acquisition parameters impact the data obtained, as discussed in Article A, valuable information can be gleaned to optimize the acquisition process when using HSI for artwork examination. This optimization is essential for ensuring the fidelity of the acquired data and enhancing the overall analysis and documentation procedures. The accurate classification of pigments in artwork materials is of great significance for conservators in determining the composition and finding their historical value. "Article B" and "Article C" explore the application of various supervised algorithms and ML models for this purpose. These articles aim to evaluate and compare different algorithms and models to identify the most suitable ones based on the specific characteristics of the materials being analyzed. This selection process is crucial to achieve an accurate and reliable pigment classification, thereby assisting conservators in their precise analysis and preservation efforts.

Moreover, "Article C" introduces the analysis of relief, which not only aids in the

investigation of artworks, but also extends its applicability to the analysis of three-dimensional objects such as sculptures, and more. This expanded scope of analysis allows for a comprehensive understanding of artistic creation, enabling deeper insights into the techniques and materials used. Many valuable CH artifacts, such as historical collections and documents, exist exclusively in microfiche format. Preserving and ensuring access to these materials has become increasingly crucial, increasing their popularity. In response to this emerging trend, the research conducted in Article D and Article E explores alternative digitization technologies and the investigation of image processing techniques to improve the quality of digitized microfiche. This research work has immense potential to propel the digitization process for microfiche materials, thus making a substantial contribution to CH.

## 4.5 Contribution to Other Application Domains

Although the primary focus of this research lies within the domain of CH imaging, the results and analyses obtained from this study have broader implications that can be extended to various other domains, allowing new perspectives and relevance across various applications.

The analysis presented in Article A focuses on the acquisition parameters for the close-range HSI within a laboratory setup for the CH application. The results of this article can be extended to other close-range applications, including forensic investigations (Edelman et al. 2012), medical imaging (Zhi et al. 2007), and food analysis (Melit Devassy et al. 2020), among others. For example, similar to CH artifacts, the duration of exposure to the sample is equally crucial in microbiology. When acquiring HSI data for live microbial samples or food analysis, prolonged exposure can cause alterations or damage to the sample (Gowen et al. 2015), necessitating careful consideration of the acquisition parameters. The detection and identification of forensic traces play a important role in crime scene investigations. HSI, in close range, has emerged as a valuable technique for the detection and visualization of latent traces using the spectral differences between the trace and its background data (Edelman et al. 2012). In addition to the significance of achieving an optimal acquisition setup, the findings and analyses presented in Article B and Article C, focusing on a range of supervised algorithms and ML models, might provide considerable potential to enhance forensic applications. Another example of an application where the results of these articles can be significant is within the field of geological imaging with HSI. In this application, valuable surface composition information is obtained by statistical comparison of the known field (that is, library spectra) with unknown image spectra (Van der Meer 2006) Likewise, it can be equally important for other applications where the effectiveness of spectral similarity measures is crucial.

The results and analysis of Article C can be useful in scenarios where HSI is employed to assess the composition and properties of materials exhibiting relief or non-flat surfaces within a close range. An illustrative example of this is the application of

HSI in the pharmaceutical industry, where it is used to identify active compounds or ingredients present in tablets or drugs (SPECIM 2020). Article D and E focus on the CH domain, particularly for microfiche digitization and enhancement. However, the findings and analysis presented in these articles are equally important for numerous other applications that involve digital imaging and magnification techniques. One such example is digital pathology (Pinco et al. 2009), where digital imaging with some image enhancement techniques allows pathologists to visualize and analyze tissue samples at a higher level of detail. Similarly, within the field of medical applications such as dermatology (Hibler et al. 2016; Pradhan et al. 2020), dermatologists use magnification, or dermoscopy (10x - 40x magnification), to improve diagnostics when assessing the skin under visible light that the naked eye cannot detect. These are just a few examples, and potential applications could extend to many other fields where detailed examination and analysis of visual data is essential.

# Chapter 5

# Conclusion and Future Directions

In this chapter, we provide a conclusive summary of the thesis and give perspective for future work on the discussed topic.

## 5.1 Conclusion

Digitization of CH artifacts is essential for improved accessibility, conservation, and research. It enables wider public access and safeguards against physical deterioration and potential loss caused by natural disasters or human actions. With the growing adoption of non-invasive imaging techniques, such as multispectral and hyperspectral imaging, in CH, digitization has increased exponentially and extended its applications. Pigment classification in CH plays a crucial role in understanding the historical context, authenticity, and conservation of artworks. HSI has become an invaluable tool in this regard, enabling detailed analysis and accurate pigment identification by capturing a wide range of spectral information. However, it also poses various challenges in terms of data quality. Likewise, microfiche, a form of microform, was extensively used for archiving and preserving historical documents prior to the emergence of digital formats; the presence of numerous historical collections and documents solely in microfiche format transforms them into invaluable CH artifacts. However, despite their long lifespan, microfiche is vulnerable to damage and requires digitization for preservation and broader accessibility. Although microfiche relies on specialized readers for direct human eye reading, available in select archives or libraries, these devices can face various challenges that affect their usability over time. Therefore, finding solutions to enhance the reading experience and facilitate the efficient utilization of microfiche resources becomes crucial in preserving and accessing these valuable materials.

This Ph.D. work has investigated and identified crucial parameters for characterizing and evaluating the behavior and content of digitized artifacts in CH. Research in this dissertation covers both conventional RGB and HSI techniques, with a specific focus on

two important CH artifacts: paintings and microfiche. This research has significantly contributed to our understanding of the importance of HSI in pigment classification. In conclusion, this research provides valuable insights into the importance of HSI for pigment classification. The study extensively investigated various aspects of HSI, including the impact of acquisition parameters on the datacube, a detailed analysis of classification algorithms using supervised and machine learning techniques for pigment identification, and the significance of data augmentations.

In addition, the research delved into the relevance of relief in artwork and its implications for pigment analysis. Furthermore, this work sheds light on the crucial aspects of microfiche digitization, emphasizing the need for advanced imaging technologies to ensure its successful implementation. The objective and subjective evaluation carried out in this study provided valuable information on the effectiveness of these technologies in preserving and improving the quality of microfiche content. This research contributes to the preservation and accessibility of CH by addressing the challenges associated with integrating quality assessment into the digitization process, particularly concerning technology adaptation. Furthermore, the knowledge and methodologies developed in this research have broader applications beyond the domain of cultural heritage. These findings can be applied to various other application domains, expanding the scope and making valuable contributions to the field of applied computer vision and image processing.

## 5.2 Future Perspective

Throughout this Ph.D. research, significant findings and insights have been obtained, leading to the identification of various research challenges and limitations, many of which have been highlighted in Chapter 4. This section will outline and discuss some of the research questions that exhibit considerable potential. By conducting research endeavors that address these research questions, further advancements can be made in the field.

**Real Object:** As highlighted in Section 4.3, one of the common limitations observed in Article A, B and C was the inaccessibility of a real painting to validate our results. Access to actual paintings would have significantly strengthened our research work, allowing for a more comprehensive analysis that accurately reflects real-world scenarios. In the future, incorporating real artworks with relief, a diverse range of pigments from different periods, and a varnished surface would not only facilitate exploration of the research potential but also allow one to tackle some new practical challenges; for instance, one might have specular reflection due to varnish surfaces, and the effectiveness of different illumination geometries could be explored as a potential solution. Furthermore, the thickness of the pigment layers affects their ability to absorb and reflect light. Thicker pigments can absorb more light, while thinner layers may transmit more light or scatter it differently. As pigment thickness can be a characteristic feature used for the identification and classification, understanding

how pigment thickness influences the spectral features would be significant.

**Preprocessing Datacube:** As emphasized in Article B (Section 4.1.2), a hypothesis is that the inclusion of computing spectrum derivatives as a preprocessing step could potentially enhance the performance of classification results. This approach is not new; derivative analysis is a well-established method in spectroscopy (BUTLER et al. 1970). Numerous studies (Demetriades-Shah et al. 1990; Holden et al. 1998; Tsai et al. 2002), mostly in remote sensing, have demonstrated the effectiveness of identifying derivative features that effectively distinguish between target classes. While applying the derivative of the spectrum can offer advantages, there are also limitations to be considered, such as the derivative can amplify noise in the signal (Tsai et al. 2002), selection of methods for computing derivative, etc. In addition, the effectiveness of the derivative of the spectrum may vary depending on the specific classification task and the characteristics of the data(Kalluri et al. 2010). Additionally, the computational complexity and additional processing time required for computing derivatives should be considered, as it may impact the practicality and real-time applicability of this approach. Therefore, further investigation is needed to determine the extent to which the addition of derivative features improves classification accuracy and whether it is applicable to a wide range of datasets and classification tasks.

As mentioned in Article C, the influence of shadows on the performance of pigment classification accuracy. Applying shadow removal techniques prior to classification and analyzing the resulting improvements would be a valuable area of investigation. Many shadow identification and removal techniques exist in remote sensing applications (X. Liu, H. Wang, et al. 2019; Omruuzun et al. 2015). However, shadow identification can be a challenging task in paintings; unlike in remote-sensing images, where shadows may have more apparent outlines and shapes(X. Liu, Hou, et al. 2017), in paintings, artistic interpretation can make it challenging to identify and interpret the shadows accurately. Therefore, future research can focus on developing techniques and methodologies capable of identifying and correcting shadows within the context of paintings.

**Multi-Device Analysis:** Although the results and analysis presented in this thesis are derived from data obtained from a specific device, it is essential to recognize that generalizing these findings requires the consideration of other HSI systems. A more comprehensive understanding of the phenomena under investigation can be achieved by including a wider range of devices, such as those employing different optics, imaging modes (e.g., whiskbroom, snapshot), and sensor technologies. (Pillay, Picollo, et al. 2020) conducted a round-robin test using data obtained from various HSI systems following routine acquisition workflow to gain valuable insights into the accuracy, reproducibility, and precision of the obtained data. The results underscore the necessity of standardized workflows in HSI. Therefore, establishing standard protocols and benchmarks for data collection, processing, and analysis would ensure reproducibility and facilitate further advances in the field.

**Objective Image Quality Metrics:** As highlighted in Article E, subjective assessment typically requires a significant amount of time, effort, and resources to conduct experiments involving human observers. This can be challenging, especially when dealing with large datasets. On the contrary, objective assessment methods can be automated, providing faster and more efficient evaluations. Considering the availability of numerous objective image quality metrics, some specifically designed for particular applications (Pedersen, Bonnier, et al. 2011) and others for general image quality (Hore et al. 2010), it is worth examining their relevance in microfiche digitization. Exploring the potential benefits involves investigating whether these metrics can be directly utilized or modified to better align with the unique characteristics of microfiche digitization.

# References

American National Standards Institute (ANSI IT9.2-1988) (1988). *American National Standard for Imaging Media-Photographic Processed Films, Plates, and Papers-Filing Enclosures and Storage Containers*. URL: https://law.resource.org/pub/us/cfr/ibr/001/aimm.it9.2.1998.pdf (visited on 01/16/2022).

Auger, C. (1991). "The Importance of Microforms." In: *Microform and Imaging Review* 20.4. Reprinted in 2009 as an online publication, Accessed: March 20, 2023, pp. 171–175. URL: https://doi.org/10.1515/mfir.1991.20.4.171.

Bai, D., D. W. Messinger, and D. Howell (2017). "Hyperspectral analysis of cultural heritage artifacts: pigment material diversity in the Gough Map of Britain." In: *Optical Engineering* 56.8, pp. 1–11. DOI: 10.1117/1.OE.56.8.081805.

Baias, M. (2020). *PAINTING LAYERS: How Scientific Analysis Reveals the Painting Structure and Composition*. Retrieved from: https://www.mariabaias.com/post/painting-layers-structure-scientific-analysis. Accessed: 21 May 2023.

Barnett, J., S. Miller, and E. Pearce (2006). "Colour and art: A brief history of pigments." In: *Optics Laser Technology* 38.4. Colour and Design in the natural and man-made worlds, pp. 445–453. ISSN: 0030-3992. DOI: https://doi.org/10.1016/j.optlastec.2005.06.005.

Barrett, S. and D. C. Stulik (1995). "An integrated approach for the study of painting techniques." In: *Historical painting techniques, materials, and studio practice: preprints of a symposium, University of Leiden, the Netherlands, 26-29 June 1995*, pp. 6–11.

Baumgartner, A., P. Gege, C. Köhler, K. Lenhard, and T. Schwarzmaier (2012). "Characterisation methods for the hyperspectral sensor HySpex at DLR's calibration home base." In: *Sensors, Systems, and Next-Generation Satellites XVI*. Vol. 8533. SPIE Remote Sensing, pp. 371–378.

Baxter, W., J. Wendt, and M. C. Lin (2004). "IMPaSTo: A Realistic, Interactive Model for Paint." In: *Proceedings of the 3rd International Symposium on Non-photorealistic*

*Animation and Rendering*, pp. 45–148. ISBN: 1581138873. DOI: 10.1145/987657.987665.

Blake, J. (2000). "On defining the cultural heritage." In: *International & Comparative Law Quarterly* 49.1, pp. 61–85.

Bland, J. M. and D. G. Altman (1995). "Multiple significance tests: the Bonferroni method." In: *British Medical Journal(BMJ)* 310.6973, p. 170.

Boust, C. and Y. Arteaga (2023). "Color and gloss measurements in cultural heritage conservation science: recent advances in France." In: *London Imaging Meeting 2023*. Society for Imaging Science and Technology (IS&T).

Boyat, A. K. and B. K. Joshi (2015). "A review paper: noise models in digital image processing." In: *arXiv preprint arXiv:1505.03489*.

British Library, N. P. O. (1992). *Microforms in Libraries: the Untapped Resource? :Papers given at the National Preservation Office Conference held 13-15 October 1992 in Birmingham*. Retrieved from: https://bl.iro.bl.uk/downloads/3783a9e9-bb14-4f4e-9a0d-0bf2bf00a8a7?locale=en. London, UK: National Preservation Office. ISBN: 0712303375.

Brooke, G. J. and H. K. Bond (1999). *Allegro Qumran Photograph Collection*. Leiden, The Netherlands: Brill. ISBN: 978-90-04-19625-4.

Brown, H., J. Baker, W. Cybulski, A. Fenton, J. Glover, P. Negus, and J. Palm (2012). "The role of microfilm in digital preservation." In: *Microform & Digitization Review* 41.2, pp. 65–82.

BT.500-14, RECOMMENDATION ITU-R (2020). *Methodologies for the subjective assessment of the quality of television images*. Accessed on June 1, 2023. Geneva, Switzerland.

BUTLER, W. t. and D. Hopkins (1970). "Higher derivative analysis of complex absorption spectra." In: *Photochemistry and Photobiology* 12.6, pp. 439–450.

Carvalho Júnior, O. de, R. Guimarães, A. Gillespie, N. Silva, and R. Gomes (2011). "A New Approach to Change Vector Analysis Using Distance and Similarity Measures." In: *Remote Sensing* 3, pp. 2473–2493. DOI: 10.3390/rs3112473.

Cavaleri, T., A. Giovagnoli, and M. Nervo (2013). "Pigments and Mixtures Identification by Visible Reflectance Spectroscopy." In: *Procedia Chemistry* 8, pp. 45–54. DOI: 10.1016/j.proche.2013.03.007.

Chein-I Chang (2000). "An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis." In: *IEEE Transactions on Information Theory* 46.5, pp. 1927–1932. DOI: 10.1109/18.857802.

Chen, M.-Y., Y.-B. Huang, S. Chang, and M. Ouhyoung (2019). "Prediction Model for Semitransparent Watercolor Pigment Mixtures Using Deep Learning with a Dataset of Transmittance and Reflectance." In: *arXiv preprint arXiv:1904.00275*.

Conway, P. (2011). "Archival quality and long-term preservation: a research framework for validating the usefulness of digital surrogates." In: *Archival Science* 11, pp. 293–309.

Corona, L. (2023). "Digitization for the visibility of collections." In: *Collection and Curation* 42.3, pp. 73–80. DOI: https://doi.org/10.1108/CC-06-2022-0024.

Cosentino, A. (2014a). "FORS Spectral Database of Historical Pigments in Different Binders." In: *E-conservation Journal* 2, pp. 57–68. DOI: 10.18236/econs2.201410.

Cosentino, A. (2014b). "FORS spectral database of historical pigments in different binders." In: *E Conservation Journal* 2, pp. 54–65. ISSN: 2189-1335.

Cucci, C., J. K. Delaney, and M. Picollo (2016). "Reflectance hyperspectral imaging for investigation of works of art: old master paintings and illuminated manuscripts." In: *Accounts of chemical research* 49.10, pp. 2070–2079.

Davenport, W. B., W. L. Root, et al. (1958). *An introduction to the theory of random signals and noise*. Vol. 159. McGraw-Hill New York.

David, H. A. (1988). *The method of paired comparisons*. 2nd ed. Charles Griffin ; Oxford University Press, London. ISBN: 9780852642900.

De Carvalho, O. A. and P. R. Meneses (2000). "Spectral correlation mapper (SCM): an improvement on the spectral angle mapper (SAM)." In: *Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18*. Vol. 9. JPL Publication Pasadena, CA.

De Haas, W. (1958). "The microfiche." In: *Journal of the American Society for Information Science* 9.2, p. 99. DOI: https://doi.org/10.1002/asi.5090090204.

Deborah, H. and D. J. Mandal (2021). "Evaluation of Text Legibility in Alternative Imaging Approaches to Microfiche Digitization." In: *Proc. IS&T Archiving*, pp. 96–101. DOI: 10.2352/issn.2168-3204.2021.1.0.22.

Deepthi, D., B. M. Devassy, S. George, P. Nussbaum, and T. Thomas (2022). "Classification of forensic hyperspectral paper data using hybrid spectral similarity algorithms." In: *Journal of Chemometrics* 36.1, e3387.

Delaney, J. K., M. Thoury, J. G. Zeibel, P. Ricciardi, K. M. Morales, and K. A. Dooley (2016). "Visible and infrared imaging spectroscopy of paintings and improved reflectography." In: *Heritage Science* 4, pp. 1–10.

Demetriades-Shah, T. H., M. D. Steven, and J. A. Clark (1990). "High resolution derivative spectra in remote sensing." In: *Remote Sensing of Environment* 33.1, pp. 55–64.

Derrien, H. (1993). "MARC, a new methodology for art reproduction in colour." In: *Information services & use* 13.4, pp. 357–369.

DiGiuseppe, M. (2011). "10.4 Electromagnetic Radiation." In: *Nelson Physics 11: University Preparation*. Nelson Education, pp. 526–531. ISBN: 9780176510374. URL: https://books.google.no/books?id=RJU6nQAACAAJ.

Digney-Peer, S., K. Thomas, R. Perry, J. Townsend, and S. Gritt (2020). "The imitative retouching of easel paintings." In: *Conservation of easel paintings*. Routledge, pp. 626–653.

Dormolen, H. van (2012). *Metamorfoze Preservation Imaging Guidelines*. Retrieved from: https://digitalizalas.eu/wp-content/uploads/2020/09/Metamorfoze_Preservation_Imaging_Guidelines.pdf. Accessed: 5 June 2023.

Eckstein, B. and R. Arlen (2021). "IEEE PROJECT 4001–STANDARDS FOR CHARACTERIZATION AND CALIBRATION OF HYPERSPECTRAL IMAGING DEVICES." In: *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 44, pp. 43–47.

Edelman, G. J., E. Gaston, T. G. Van Leeuwen, P. Cullen, and M. C. Aalders (2012). "Hyperspectral imaging for non-contact analysis of forensic traces." In: *Forensic science international* 223.1-3, pp. 28–39.

Elkhuizen, W., T. Dore-Callewaert, E. Leonhardt, A. Vandivere, Y. Song, S. Pont, J. M. Geraedts, and J. Dik (2019). "Comparison of three 3D scanning techniques for paintings, as applied to Vermeer's 'Girl with a Pearl Earring'." eng. In: *Heritage science* 7.1, pp. 1–22. ISSN: 2050-7445.

Engeldrum, P. G. (2001). "Psychometric scaling: avoiding the pitfalls and hazards." In: *Image Processing, Image Quality, Image Capture Systems Conference (PICS)*, pp. 101–107.

Engeldrum, P. G. (2004). "A theory of image quality: The image quality circle." In: *Journal of imaging science and technology* 48.5, pp. 447–457.

Federal Agencies Digitization Guidelines Initiative (FADGI) (2010). *Technical Guidelines for Digitizing Cultural Heritage Materials: Creation of Raster Image Files*. Retrieved from: https://www.digitizationguidelines.gov/guidelines/FADGI_Still_Image-Tech_Guidelines_2010-08-24.pdf. Accessed: 5 June 2023.

Feller, R. L. (1986). *Artists' Pigments: A Handbook of Their History and Characteristics*. v. 1. National Gallery of Art, Washington. ISBN: 978-1-904982-74-6.

Fichner-Rathus, L. (2008). *Understanding Art: A Concise History*. Thomson Advantage Books. Thomson Wadsworth. ISBN: 9780495101680. URL: https://books.google.no/books?id=RC2GAAAACAAJ.

Fitzpatrick, R. (2008). *Maxwell's Equations and the Principles of Electromagnetism*. G - Reference,Information and Interdisciplinary Subjects Series. Infinity Science Press LLC. ISBN: 9781934015209. URL: https://books.google.no/books?id=4QaVSxxnyWwC.

Foster, D. and K. Amano (2019). "Hyperspectral imaging in color vision research: Tutorial." In: *Journal of the Optical Society of America A* 36, p. 606. DOI: 10.1364/JOSAA.36.000606.

Fryskowska, A. and J. Stachelek (2018). "A no-reference method of geometric content quality analysis of 3D models generated from laser scanning point clouds for hBIM." In: *Journal of Cultural Heritage* 34. Technoheritage 2017, pp. 95–108. ISSN: 1296-2074. DOI: https://doi.org/10.1016/j.culher.2018.04.003.

Fu, Y., H. Yu, C.-K. Yeh, J. Zhang, and T.-Y. Lee (2018). "High relief from brush painting." In: *IEEE transactions on visualization and computer graphics* 25.9, pp. 2763–2776.

Gao, X., W. Lu, D. Tao, and X. Li (2009). "Image quality assessment based on multiscale geometric analysis." In: *IEEE Transactions on Image Processing* 18.7, pp. 1409–1423.

Garcia-Spitz, C. (2017). "Patrolling the Past: Bringing the Papua New Guinea Colonial-Era Reports into the Digital Realm." In: *UC San Diego: Library*.

Gauglitz, G. and J. P. Dakin (2006). "Spectroscopic analysis." In: *Handbook of Optoelectronics Enabling Technologies (Volume II)*. Ed. by J. P. Dakin and R. G. W. Brown. Boca Raton: CRC Press, Taylor  Francis Group, pp. 1399–1442.

George, S., J. Y. Hardeberg, J. Linhares, L. Macdonald, C. Montagner, S. Nascimento, M. Picollo, R. Pillay, T. Vitorino, and E. K. Webb (2019). "A Study of Spectral Imaging Acquisition and Processing for Cultural Heritage." eng. In: *Digital Techniques for Documenting and Preserving Cultural Heritage*. Amsterdam: Amsterdam University Press, pp. 141–158. ISBN: 1942401353.

Gilroy, D., I. M. Godfrey, and I. Loo (2017). "Paintings." In: *Conservation and Care of Collections*. Western Australian Museum's Department of Materials Conservation (The original text was compiled in 1998). URL: https://manual.museum.wa.gov.au/book/export/html/69.

Goetze, J. P., F. Anders, S. Petry, J. F. Witte, and H. Lokstein (2022). "Spectral characterization of the main pigments in the plant photosynthetic apparatus by theory and experiment." In: *Chemical Physics* 559, p. 111517.

Gonzalez, V., M. Cotte, G. Wallez, A. van Loon, W. De Nolf, M. Eveno, K. Keune, P. Noble, and J. Dik (2019). "Unraveling the Composition of Rembrandt's Impasto through the Identification of Unusual Plumbonacrite by Multimodal X-ray Diffraction Analysis." In: *Angewandte Chemie* 131.17, pp. 5675–5678.

Gowen, A. A., Y. Feng, E. Gaston, and V. Valdramidis (2015). "Recent applications of hyperspectral imaging in microbiology." In: *Talanta* 137, pp. 43–54.

Groen, K. (1997). "Investigation of the use of the binding medium by Rembrandt." In: *Zeitschrift fur Kunsttechnologie und Konservierung* 2.2, pp. 208–211.

Gu, K., G. Zhai, X. Yang, W. Zhang, and M. Liu (2013). "Subjective and objective quality assessment for images with contrast change." In: *2013 IEEE International Conference on Image Processing*, pp. 383–387. DOI: 10.1109/ICIP.2013.6738079.

Gueli, A. M., G. Bonfiglio, S. Pasquale, and S. O. Troja (2017). "Effect of particle size on pigments colour." In: *Color Research & Application* 42.2, pp. 236–243.

Hardeberg, J. Y. (2001). *Acquisition and Reproduction of Color Images: Colorimetric and Multispectral Approaches*. Universal-Publishers. ISBN: 1581121350.

Hardeberg, J. Y., F. Schmitt, and H. Brettel (2002). "Multispectral color image capture using a liquid crystal tunable filter." In: *Optical engineering* 41.10, pp. 2532–2548.

Harley, R. D. ( D. (1982). *Artists' pigments c. 1600-1835 : a study in English documentary sources*. 2nd ed. Technical studies in the arts, archaeology, and architecture. London ; Boston: Butterworth Scientific. ISBN: 0408709456.

Hawken, W. R. (1975a). *Evaluating Microfiche Readers: A Handbook for Librarians*. Washington: Council on Library Resources.

Hawken, W. R. (1975b). *Evaluating Microfiche Readers: A Handbook for Librarians*. en. Google-Books-ID: KfVCAAAAIAAJ. Council on Library Resources.

Hayem-Ghez, A., C. Boust, E. Ravaud, G. Bastian, M. Menu, and N. Brodie-Linder (2015). "Comparing spectrophotometry and photography with hyperspectral imaging for pigments' characterization on paintings." In: *Color and Imaging Conference*. Vol. 2015. 1. Society for Imaging Science and Technology, pp. 180–185.

Hibler, B. P., Q. Qi, and A. M. Rossi (2016). "Current state of imaging in dermatology." In: *Semin Cutan Med Surg* 35.1, pp. 2–8.

Hirtle, P. B. (2002). "The impact of digitization on special collections in libraries." In: *Libraries & Culture*, pp. 42–52.

Holden, H. and E. LeDrew (1998). "Spectral discrimination of healthy and non-healthy corals based on cluster analysis, principal components analysis, and derivative spectroscopy." In: *Remote sensing of environment* 65.2, pp. 217–224.

Hore, A. and D. Ziou (2010). "Image quality metrics: PSNR vs. SSIM." In: *2010 20th international conference on pattern recognition*. IEEE, pp. 2366–2369.

Horik, R. V. (2005). "Permanent pixels:Building blocks for the longevity of digital surrogates of historical photographs." PhD thesis. Netherlands: Delft University of Technology.

Hoyer, R. W., B. B. Hoyer, P. B. Crosby, W. E. Deming, et al. (2001). "What is Quality?" In: *Quality progress* 34.7, pp. 53–62.

Hu, W., Y. Huang, W. Li, F. Zhang, and H. Li (2015). "Deep Convolutional Neural Networks for Hyperspectral Image Classification." In: *Journal of Sensors* 2015, pp. 1–12. DOI: 10.1155/2015/258619.

IEEE Geoscience and Remote Sensing Society (2018). *P4001: Standard for Characterization and Calibration of Ultraviolet through Shortwave Infrared (250 nm to 2500 nm) Hyperspectral Imaging Devices*. URL: https://standards.ieee.org/ieee/4001/7314/ (visited on 01/25/2023).

Image Access GmbH, Germany (2023). *Quality controlled scanning*. URL: https://support.imageaccess.de/downloads/product_manuals/FAQ/FAQ-Quality-Controlled-Scanning.pdf (visited on 06/05/2023).

ISO 19264-1:2021 (2021). *Photography – Archiving systems —- Imaging systems quality analysis – Part 1: Reflective originals*. Accessed on June 1, 2023. Geneva, Switzerland.

ISO 20462-1:2005 (2005). *Photography – Psychophysical experimental methods for estimating image quality —- Part 1:Overview of psychophysical elements*. Accessed on June 1, 2023. Geneva, Switzerland.

ISO 9000:2015 (2015). *Quality management systems—Fundamentals and vocabulary*. Accessed on May 7, 2023. Geneva, Switzerland.

Jagalingam, P. and A. V. Hegde (2015). "A review of quality metrics for fused image." In: *Aquatic Procedia* 4, pp. 133–142.

Janssen, R. (2001). *Computational Image Quality*. Press Monographs. Society of Photo Optical. ISBN: 9780819441324.

Jeffers, C. (2015). "The ethics and politics of cultural preservation." In: *The Journal of Value Inquiry* 49, pp. 205–220.

Kadyrova, A., M. Pedersen, B. Ahmad, D. J. Mandal, M. Nguyen, and P. H. Zimmermann (2022). "Image enhancement dataset for evaluation of image quality metrics." In: *Electronic Imaging* 34, pp. 1–6.

Kale, K. V., M. M. Solankar, D. B. Nalawade, R. K. Dhumal, and H. R. Gite (2017). "A Research Review on Hyperspectral Data Processing and Analysis Algorithms." In: *Proceedings of the national academy of sciences, India section a: physical sciences* 87, pp. 541–555.

Kalluri, H. R., S. Prasad, and L. M. Bruce (2010). "Decision-level fusion of spectral reflectance and derivative information for robust hyperspectral land cover classification." In: *IEEE Transactions on Geoscience and Remote Sensing* 48.11, pp. 4047–4058.

Kamble, V. and K. Bhurchandi (2015). "No-reference image quality assessment algorithms: A survey." In: *Optik* 126.11-12, pp. 1090–1097.

Keelan, B. W. (2002). *Handbook of image quality: characterization and prediction*. CRC Press.

Kenney, A. R. and O. Y. Rieger (2000). *Moving Theory into Practice: Digital Imaging for Libraries and Archives*. New York, NY: Research Libraries Group, p. 189.

Kerekes, J. P. and S. M. Hsu (2004). "Spectral quality metrics for VNIR and SWIR hyperspectral imagery." In: *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X*. Vol. 5425. SPIE, pp. 549–557.

Khan, M. J., H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas (2018). "Modern trends in hyperspectral image analysis: A review." In: *Ieee Access* 6, pp. 14118–14129.

Kruse, F., A. Lefkoff, J. Boardman, K. Heidebrecht, A. Shapiro, P. Barloon, and A. Goetz (1993). "The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data." In: *Remote Sensing of Environment* 44.2. Airbone Imaging Spectrometry, pp. 145–163. ISSN: 0034-4257. DOI: https://doi.org/10.1016/0034-4257(93)90013-N.

Kubik, M. (2007). "Chapter 5 Hyperspectral Imaging: A New Technique for the Non-Invasive Study of Artworks." In: *Physical Techniques in the study of Art, Archaeology and Cultural Heritage*. Ed. by D. Creagh and D. Bradley. Vol. 2. Elsevier, pp. 199–259. DOI: https://doi.org/10.1016/S1871-1731(07)80007-8.

Lee, S. D. (2001). *Digital Imaging: A Practical Handbook*. New York, NY: Neal Schuman Pub, p. 194.

Levenshtein, V. I. (1966). "Binary Codes Capable of Correcting Deletions, Insertions and Reversals." In: *Soviet Physics Doklady* 10, p. 707. URL: http://adsabs.harvard.edu/abs/1966SPhD...10..707L (visited on 03/29/2021).

Li, H., W. S. Lee, K. Wang, R. Ehsani, and C. Yang (2014). "Extended spectral angle mapping (ESAM)'for citrus greening disease detection using airborne hyperspectral imaging." In: *Precision Agriculture* 15.2, pp. 162–183.

Li, Q., X. He, Y. Wang, H. Liu, D. Xu, and F. Guo (2013). "Review of spectral imaging technology in biomedical engineering: achievements and challenges." In: *Journal of biomedical optics* 18.10, pp. 1–28.

Liang, H., D. Saunders, J. Cupitt, and C. Lahanier (2004). "1Multispectral Imaging for Easel and Wall Paintings." In: *Conservation of Ancient Sites on the Silk Road: Proceedings of the Second International Conference on the Conservation of Grotto Sites*. Ed. by N. Agnew. Getty Publications, pp. 267–274.

Linhares, J. M., J. A. Monteiro, A. Bailão, L. Cardeira, T. Kondo, S. Nakauchi, M. Picollo, C. Cucci, A. Casini, L. Stefani, et al. (2020). "How Good Are RGB Cameras Retrieving Colors of Natural Scenes and Paintings?—A Study Based on Hyperspectral Imaging." In: *Sensors* 20.21, pp. 1–14.

Liu, L., T. Miteva, G. Delnevo, S. Mirri, P. Walter, L. de Viguerie, and E. Pouyet (2023). "Neural Networks for Hyperspectral Imaging of Historical Paintings: A Practical Review." In: *Sensors* 23.5, p. 2419.

Liu, N., Y. Guo, H. Jiang, and W. Yi (2020). "Gastric cancer diagnosis using hyperspectral imaging with principal component analysis and spectral angle mapper." In: *Journal of Biomedical Optics* 25.6, p. 066005.

Liu, X., Z. Hou, Z. Shi, Y. Bo, and J. Cheng (2017). "A shadow identification method using vegetation indices derived from hyperspectral data." In: *International Journal of Remote Sensing* 38.19, pp. 5357–5373.

Liu, X., H. Wang, Y. Meng, and M. Fu (2019). "Classification of hyperspectral image by CNN based on shadow area enhancement through dynamic stochastic resonance." In: *IEEE Access* 7, pp. 134862–134870.

Lyu, S., Y. Liu, M. Hou, Q. Yin, W. Wu, and X. Yang (2020). "Quantitative analysis of mixed pigments for Chinese paintings using the improved method of ratio spectra derivative spectrophotometry based on mode." In: *Heritage Science* 8.1, pp. 1–21.

MacBeth, R. (2012). "The Technical Examination and Documentation of Easel Paintings." In: *Conservation of Easel Paintings*. Ed. by J. H. Stoner and R. Rushfield. Routledge, New York, pp. 291–305.

MacDonald, L. (2006). *Digital Heritage : Applying Digital Imaging to Cultural Heritage*. London: Butterworth-Heinemann, p. 592.

Mandal, D. J., S. George, M. Pedersen, and C. Boust (2021). "Influence of Acquisition Parameters on Pigment Classification using Hyperspectral Imaging." In: *Journal of Imaging Science and Technology* 65.5, 050406-1–050406–13.

Mandal, D. J., S. George, M. Pedersen, and C. Boust (2023). "Quality Evaluation in Cultural Heritage Digitization." In: *Cultural Heritage Analysis for new Generations (CHANGE)*. Submitted for publication. Routledge.

Mandal, D. J., H. Deborah, and M. Pedersen (2023). "Subjective Quality Evaluation of Alternative Imaging Techniques for Microfiche Digitization." In: *Journal of Cultural Heritage*. Accepted for publication.

Mandal, D. J., M. Pedersen, S. George, and C. Boust (2023). "Comparison of Pigment Classification Algorithms on Non–Flat Surfaces using Hyperspectral Imaging." In: Manuscript under revision in Journal.

Mandal, D. J., M. Pedersen, S. George, H. Deborah, and C. Boust (2023). "An Experiment–based Comparative Analysis of Pigment Classification Algorithms using Hyperspectral Imaging." In: *Journal of Imaging Science and Technology* 67.3, pp. 030403-1–030403-18. DOI: 10.2352/J.ImagingSci.Technol.2023.67.3.030403.

Mantiuk, R. K., A. Tomaszewska, and R. Mantiuk (2012). "Comparison of four subjective methods for image quality assessment." In: *Computer Graphics Forum* 31.8, pp. 2478–2491.

Martinez, K. (1991). "High resolution digital imaging of paintings: the vasari project." In: *Microcomputers for Information Management* 8.4, pp. 277–283.

Martinez, K., J. Cupitt, D. Saunders, and R. Pillay (2002). "Ten years of art imaging research." In: *Proceedings of the IEEE* 90.1, pp. 28–41.

Martinez, K. and A. Hamber (1989). "Towards A Colorimetric Digital Image Archive For The Visual Arts." In: *Electronic Imaging Applications in Graphic Arts*. Ed. by K. S. Cloud. Vol. 1073. International Society for Optics and Photonics. SPIE, pp. 114–121. DOI: 10.1117/12.952562.

Mayer, R. (1966). *The Painter's Craft: An Introduction to Artists' Methods and Materials*. Studio book. Van Nostrand. URL: https://books.google.no/books?id=wOpXAAAAYAAJ.

Melgani, F. and L. Bruzzone (2004). "Classification of hyperspectral remote sensing images with support vector machines." In: *IEEE Transactions on geoscience and remote sensing* 42.8, pp. 1778–1790.

Melit Devassy, B. and S. George (2020). "Contactless classification of strawberry using hyperspectral imaging." In: *CEUR Workshop Proceedings*.

Mishra, P., M. S. M. Asaari, A. Herrero-Langreo, S. Lohumi, B. Diezma, and P. Scheunders (2017). "Close range hyperspectral imaging of plants: A review." In: *Biosystems Engineering* 164, pp. 49–67.

National Institute of Standards and Technology (NIST) (2015). *Hyperspectral Imaging Standards*. URL: https://www.nist.gov/programs-projects/hyperspectral-imaging-standards (visited on 01/25/2023).

Ngo, K. V., J. J. Storvik, C. A. Dokkeberg, I. Farup, and M. Pedersen (2015). "QuickEval: A web application for psychometric scaling experiments." In: *Image Quality and System Performance XII*. Ed. by M.-C. Larabi and S. Triantaphillidou. SPIE, 93960O. DOI: 10.1117/12.2077548.

Norsk Elektro Optikk (2023). *HySpex SWIR-384*. URL: https://www.hyspex.com/hyspex-products/hyspex-classic/hyspex-swir-384/ (visited on 05/20/2023).

Norsk Elektro Optikk (2023). *HySpex VNIR-1800*. URL: https://www.hyspex.com/hyspex-products/hyspex-classic/hyspex-vnir-1800/ (visited on 05/20/2023).

Northeast Document Conservation Center (NEDCC) (2017). *6.1 Microfilm and Microfiche*. Northeast Document Conservation Center.

Omruuzun, F., D. O. Baskurt, H. Daglayan, and Y. Y. Cetin (2015). "Shadow removal from VNIR hyperspectral remote sensing imagery with endmember signature analysis." In: *Next-Generation Spectroscopic Technologies VIII*. Vol. 9482. SPIE, pp. 360–367.

Oregon State Archives (2013). *Micrographics Manual*. Retrieved from: https://sos.oregon.gov/archives/Documents/recordsmgmt/scd/micrographics.pdf. Accessed: 16 May 2023.

P.910, Recommendation ITU-T (2022). *Subjective video quality assessment methods for multimedia applications*. Accessed on June 1, 2023. Geneva, Switzerland.

Park, B., W. Windham, K. Lawrence, and D. Smith (2007). "Contaminant Classification of Poultry Hyperspectral Imagery using a Spectral Angle Mapper Algorithm." In: *Biosystems Engineering* 96.3, pp. 323–333. ISSN: 1537-5110. DOI: https://doi.org/10.1016/j.biosystemseng.2006.11.012.

Park, J.-I., M.-H. Lee, M. D. Grossberg, and S. K. Nayar (2007). "Multispectral imaging using multiplexed illumination." In: *2007 IEEE 11th International Conference on Computer Vision*. IEEE, pp. 1–8.

Pedersen, M. (2011). "Image quality metrics for the evaluation of printing workflows." PhD thesis. University of Oslo.

Pedersen, M., N. Bonnier, J. Y. Hardeberg, and F. Albregtsen (2011). "Image quality metrics for the evaluation of print quality." In: *Image quality and system performance VIII*. Vol. 7867. SPIE, pp. 11–29.

Pedersen, M., N. Bonnier, J. Y. Hardeberg, and F. Albregtsen (2010). "Attributes of image quality for color prints." In: *Journal of Electronic Imaging* 19.1, pp. 011016–011016.

Pedersen, M. and J. Y. Hardeberg (2012). "Full-reference image quality metrics: Classification and evaluation." In: *Foundations and Trends® in Computer Graphics and Vision* 7.1, pp. 1–80.

Phillips, J. B. and H. Eliasson (2018). *Camera image quality benchmarking*. John Wiley & Sons.

Picollo, M., C. Cucci, A. Casini, and L. Stefani (2020). "Hyper-Spectral Imaging Technique in the Cultural Heritage Field: New Possible Scenarios." In: *Sensors* 20.10. ISSN: 1424-8220. DOI: 10.3390/s20102843.

Pillay, R., J. Y. Hardeberg, and S. George (2019). "Hyperspectral imaging of art: Acquisition and calibration workflows." In: *Journal of the American Institute for Conservation* 58.1-2, pp. 3–15.

Pillay, R., M. Picollo, J. Y. Hardeberg, and S. George (2020). "Evaluation of the Data Quality from a Round-Robin Test of Hyperspectral Imaging Systems." In: *Sensors* 20.14, p. 3812.

Pinco, J., R. A. Goulart, C. N. Otis, J. Garb, and L. Pantanowitz (2009). "Impact of digital image manipulation in cytology." In: *Archives of Pathology & Laboratory Medicine* 133.1, pp. 57–61.

Pinson, M. H. and S. Wolf (2003). "Comparing subjective video quality testing methodologies." In: *Visual Communications and Image Processing 2003*. Vol. 5150. SPIE, pp. 573–582.

Plisson, J. S., L. de Viguerie, L. Tahroucht, M. Menu, and G. Ducouret (2014). "Rheology of white paints: How Van Gogh achieved his famous impasto." In: *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 458, pp. 134–141.

Pouyet, E., T. Miteva, N. Rohani, and L. de Viguerie (2021). "Artificial Intelligence for Pigment Classification Task in the Short-Wave Infrared Range." In: *Sensors* 21.18. Article No. 6150. ISSN: 1424-8220. DOI: 10.3390/s21186150.

Pradhan, S., X. Ran, S. Xue, Y. Ran, et al. (2020). "Dermoscopic manifestations of nail diseases." In: *Dermatologica Sinica* 38.4, p. 205.

Prott, L. V. and P. J. O'Keefe (1992). "'Cultural heritage'or 'cultural property'?" In: *International Journal of Cultural Property* 1.2, pp. 307–320.

Qin, J. (2010a). "CHAPTER 2 - Spectral Preprocessing and Calibration Techniques." In: *Hyperspectral imaging for food quality analysis and control*. Academic Press, pp. 45–78. ISBN: 978-0-12-374753-2. DOI: https://doi.org/10.1016/B978-0-12-374753-2.10002-4.

Qin, J. (2010b). "CHAPTER 5 - Hyperspectral Imaging Instruments." In: *Hyperspectral Imaging for Food Quality Analysis and Control*. Ed. by D.-W. Sun. San Diego: Academic Press, pp. 129–172. ISBN: 978-0-12-374753-2. DOI: https://doi.org/10.1016/B978-0-12-374753-2.10005-X.

Quye, A. and M. Strlič (2019). "Ethical Sampling Guidance." In: *Institute of Conservation: London, UK*, pp. 1–14.

Ray, S. (2011). "Chapter 6 - Photographic and geometrical optics." In: *The Manual of Photography (Tenth Edition)*. Ed. by E. Allen and S. Triantaphillidou. Tenth Edition. Oxford: Focal Press, pp. 103–117. ISBN: 978-0-240-52037-7. DOI: https://doi.org/10.1016/B978-0-240-52037-7.10006-7.

Reinhard, E., E. A. Khan, A. O. Akyuz, and G. Johnson (2008). *Color imaging: fundamentals and applications*. CRC Press. ISBN: 9781568813448.

Rohani, N., E. Pouyet, M. Walton, O. Cossairt, and A. K. Katsaggelos (2018). "Nonlinear unmixing of hyperspectral datasets for the study of painted works of art." In: *Angewandte Chemie* 130.34, pp. 11076–11080.

Rubens, P. P. (n.d.). *The Virgin as the Woman of the Apocalypse*. about 1623–1624. J. Paul Getty Museum. URL: https://www.getty.edu/art/collection/object/103RGR#full-artwork-details.

Saffady, W. et al. (1978). *Computer-output microfilm*. American Library Association.

Schreiner, M., B. Frühmann, D. Jembrih-Simbürger, and R. Linke (2004). "X-rays in art and archaeology: an overview." In: *Powder Diffraction* 19.1, pp. 3–11.

Schreiner, M., R. Wiesinger, and W. Vetter (2017). "Identification and Preservation of Cultural Heritage." In: *New York: ChemistryViews*. Accessed: April 20, 2023. URL: https://www.chemistryviews.org/details/ezine/9610631/Identification_and_Preservation_of_Cultural_Heritage/.

Selle, T. (2003). *Acetate Microfilm*. New York State Archives, Retrieved from: https://www.archives.nysed.gov/sites/archives/files/mr_pub15.pdf. Accessed: 16 May 2023.

Shaw, G. and D. Manolakis (2002). "Signal processing for hyperspectral image exploitation." In: *IEEE Signal Processing Magazine* 19.1, pp. 12–16. DOI: 10.1109/79.974715.

Shrestha, R. and J. Y. Hardeberg (2013). "Multispectral imaging using LED illumination and an RGB camera." In: *Color and Imaging Conference*. Vol. 2013. 1. Society for Imaging Science and Technology, pp. 8–13.

Shrestha, R., A. Mansouri, and J. Y. Hardeberg (2011). "Multispectral imaging using a stereo camera: Concept, design and assessment." In: *EURASIP Journal on Advances in Signal Processing* 2011.1, pp. 1–15.

Shrestha, R., R. Pillay, S. George, and J. Y. Hardeberg (2014). "Quality Evaluation in Spectral Imaging–Quality Factors and Metrics." In: *Journal of the International Colour Association* 12, pp. 22–35.

Sinclair, E. (1995). "The polychromy of Exeter and Salisbury Cathedrals: a preliminary comparison." In: *Historical painting techniques, materials, and studio practice: preprints of a symposium, University of Leiden, the Netherlands, 26-29 June 1995*. Getty Conservation Institute Malibu, pp. 105–110.

Singh, G. and A. Mittal (2014). "Various image enhancement techniques-a critical review." In: *International Journal of Innovation and Scientific Research* 10.2, pp. 267–274.

SPECIM (2020). *Hyperspectral imaging solution for the pharmaceutical and biotech industry*. URL: https://www.specim.com/hyperspectral-imaging-solution-for-the-pharmaceutical-and-biotech-industry/ (visited on 07/07/2023).

SPECIM, Spectral Imaging Ltd. (2023a). *Specim FX50*. URL: https://www.specim.com/products/specim-fx50/ (visited on 05/20/2023).

SPECIM, Spectral Imaging Ltd. (2023b). *Spectral Camera LWIR*. URL: https://qd-europe.com/fileadmin/Mediapool/products/Specim/_pdf/Spectral_cameras_LWIR.pdf (visited on 05/20/2023).

*Spectralon multi-step targets* (2023). URL: https://www.labsphere.com/product/spectralon-reflectance-targets/ (visited on 05/01/2023).

SphereOptics (2023). *Zenith Polymer® Diffuse Reflectance Standards*. URL: https://sphereoptics.de/en/product/zenith-polymer-reflectance-standards/ (visited on 05/01/2023).

Špiclin, Ž., J. Katrašnik, M. Bürmen, F. Pernuš, and B. Likar (2010). "Geometric calibration of a hyperspectral imaging system." In: *Applied optics* 49.15, pp. 2813–2818.

Spigai, F. G. (1973). *The Invisible Medium: The State of the Art of Microform and a Guide to the Literature*. ED075029. URL: http://files.eric.ed.gov/fulltext/ED075029.pdf (visited on 05/16/2023).

Stefanou, M. S. and J. P. Kerekes (2009). "A Method for Assessing Spectral Image Utility." In: *IEEE Transactions on Geoscience and Remote Sensing* 47.6, pp. 1698–1706. DOI: 10.1109/TGRS.2008.2006364.

Stoveland, L. P., M. Stols-Witlox, B. Ormsby, and N. Streeton (2021). "Mock-ups and materiality in conservation research." In: *Transcending Boundaries: Integrated Approaches to Conservation. ICOM-CC 19th Triennial Conference Preprints, Beijing. International Council of Museums, Paris*, pp. 1–14.

Stoveland, L. P., T. Frøysaker, M. Stols-Witlox, T. Grøntoft, C. C. Steindal, O. Madden, and B. Ormsby (2021). "Evaluation of novel cleaning systems on mock-ups of unvarnished oil paint and chalk-glue ground within the Munch Aula Paintings Project." In: *Heritage Science* 9, pp. 1–32.

Stuart, B. H. (2007). "Conservation Materials." In: *Analytical Techniques in Materials Conservation*. John Wiley  Sons, Ltd. Chap. 1, pp. 1–42. ISBN: 9780470060520. DOI: https://doi.org/10.1002/9780470060520.ch1.

Stulik, D., D. Miller, H. Khanjian, N. Khandekar, J. Carlson, R. Wolbers, and W. C. Petersen (2004). *Solvent gels for the cleaning of works of art: the residue question*. Getty Publications.

Sweet, J., J. Granahan, and M. Sharp (2000). "An objective standard for hyperspectral image quality." In: *Proceedings of 9th AVIRIS Earth Science and Applications Workshop*. Jet Propulsion Laboratory. Pasadena, California, p. 10.

Taft, W. S. and J. W. Mayer (2000). *The science of paintings*. Springer New York, NY. ISBN: 978-0-387-98722-4. DOI: 10.1007/b97567.

Tao, D., X. Li, W. Lu, and X. Gao (2009). "Reduced-reference IQA in contourlet domain." In: *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 39.6, pp. 1623–1627.

Thurstone, L. L. (2017). "A law of comparative judgment." In: *Scaling*. Routledge, pp. 81–92.

Tigchelaar, E. (2004). "ON THE UNIDENTIFIED FRAGMENTS OF "DJD" XXXIII AND PAM 43.680: A NEW MANUSCRIPT OF "4QNARRATIVE AND POETIC COMPO-

SITION", AND FRAGMENTS OF "4Q13", "4Q269", "4Q525" AND "4QSb"(?)" In: *Revue de Qumrân* 21.3 (83), pp. 477–485.

Tite, M. (2002). "Archaeological collections: invasive sampling versus object integrity." In: *Papers from the Institute of Archaeology* 13, pp. 1–5.

Tov, E. and S. Pfann (1995). *Companion Volume to the Dead Sea Scrolls on Microfiche Edition: Published under the Auspices of the Israel Antiquities Authority*. Leiden, The Netherlands: Brill. ISBN: 978-90-04-10288-0.

Townsend, J. H. (1995). "Painting techniques and materials of Turner and other British artists 1775-1875." In: *Historical painting techniques, materials, and studio practice: preprints of a symposium, University of Leiden, the Netherlands, 26-29 June 1995*, pp. 176–185.

Tripathi, M. K. and H. Govil (2019). "Evaluation of AVIRIS-NG hyperspectral images for mineral identification and mapping." In: *Heliyon* 5.11, e02931. ISSN: 2405-8440. DOI: https://doi.org/10.1016/j.heliyon.2019.e02931.

Tsai, F. and W. D. Philpot (2002). "A derivative-aided hyperspectral image analysis system for land-cover classification." In: *IEEE Transactions on Geoscience and Remote Sensing* 40.2, pp. 416–425.

Van der Meer, F. (2006). "The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery." In: *International journal of applied earth observation and geoinformation* 8.1, pp. 3–17.

Verhoeven, G. (2016). "Basics of photography for cultural heritage imaging." In: *3D recording, documentation and management of cultural heritage*. Ed. by E. Stylianidis and F. Remondino. Whittles Publishing, pp. 127–251.

Vishnu, S., R. R. Nidamanuri, and R. Bremananth (2013). "Spectral material mapping using hyperspectral imagery: a review of spectral matching and library search methods." In: *Geocarto international* 28.2, pp. 171–190.

Wandell, B. A. (1995). *Foundations of vision*. Sinauer Associates. ISBN: 0878938532.

Wang, L. (2005). *Support vector machines: theory and applications*. Vol. 177. Springer Science & Business Media.

Wang, Z., A. Bovik, H. Sheikh, and E. Simoncelli (2004). "Image Quality Assessment: From Error Visibility to Structural Similarity." In: *Image Processing, IEEE Transactions on* 13, pp. 600–612. DOI: 10.1109/TIP.2003.819861.

Wang, Z. and A. C. Bovik (2006). *Modern Image Quality Assessment*. Morgan & Claypool Publishers. ISBN: 9781598290226. DOI: 10.1007/978-3-031-02238-8_1.

Webb, E., S. Robson, and R. Evans (2020). "QUANTIFYING DEPTH OF FIELD AND SHARPNESS FOR IMAGE-BASED 3D RECONSTRUCTION OF HERITAGE OBJECTS." In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLIII-B2-2020, pp. 911–918. DOI: 10.5194/isprs-archives-XLIII-B2-2020-911-2020.

Weerd, J. van der, A. van Loon, and J. J. Boon (2005). "FTIR Studies of the Effects of Pigments on the Aging of Oil." In: *Studies in Conservation* 50.1, pp. 3–22. ISSN: 00393630. URL: http://www.jstor.org/stable/25487713 (visited on 01/03/2023).

Widenhorn, R., J. C. Dunlap, and E. Bodegom (2010). "Exposure Time Dependence of Dark Current in CCD Imagers." In: *IEEE Transactions on Electron Devices* 57.3, pp. 581–587. DOI: 10.1109/TED.2009.2038649.

Wu, T., G. Li, Z. Yang, H. Zhang, Y. Lei, N. Wang, and L. Zhang (2017). "Shortwave infrared imaging spectroscopy for analysis of ancient paintings." In: *Applied spectroscopy* 71.5, pp. 977–987.

Yamaguchi, M., H. Haneishi, and N. Ohyama (2008). "BeyondRred–Green–Blue (RGB): spectrum-Based Color Imaging Technology." In: *Journal of Imaging Science and Technology* 52.1, pp. 10201–15.

Yan, L., M. Zhao, X. Wang, Y. Zhang, and J. Chen (2021). "Object detection in hyperspectral images." In: *IEEE Signal Processing Letters* 28, pp. 508–512.

Yu, S., S. Jia, and C. Xu (2017). "Convolutional neural networks for hyperspectral image classification." In: *Neurocomputing* 219, pp. 88–98.

*ZECCHI* (2023). URL: https://zecchi.it/ (visited on 06/10/2023).

Zeutschel GmbH (2023a). *delta / delta plus, The Universal Film Scanning Systems*. URL: https://thietbisohoa.vn/sites/default/files/2019-04/zeutschel_delta_deltaplus_en.pdf (visited on 05/16/2023).

Zeutschel GmbH (2023b). *Microfilm scanner delta+ Advanced Microfilm Scanner*. URL: https://www.zeutschel.de/en/produkte/deltaadvanced/ (visited on 05/16/2023).

Zhi, L., D. Zhang, J.-q. Yan, Q.-L. Li, and Q.-l. Tang (2007). "Classification of hyperspectral medical tongue images for tongue diagnosis." In: *Computerized Medical Imaging and Graphics* 31.8, pp. 672–678.

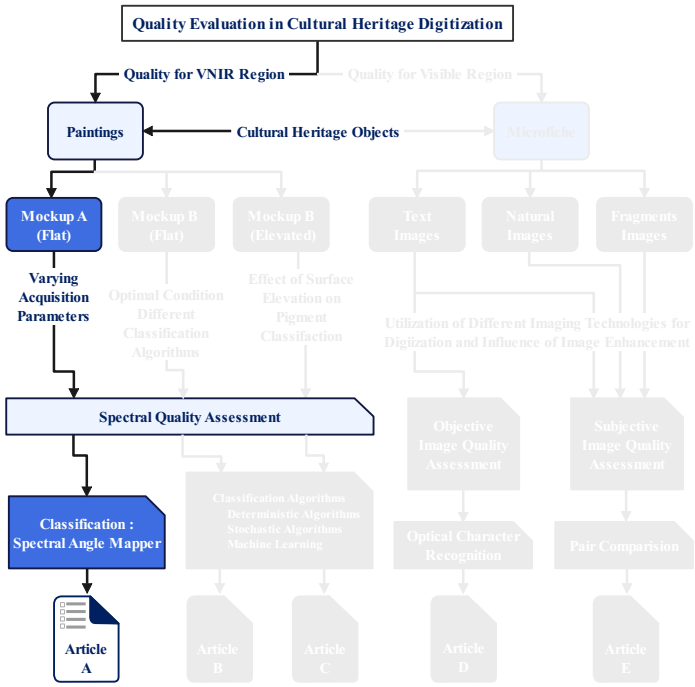
# Part II

# Original Articles

# Article A

## Influence of Acquisition Parameters on Pigment Classification using Hyperspectral Imaging

D. J. Mandal, S. George, M. Pedersen, and C. Boust (2021). “Influence of Acquisition Parameters on Pigment Classification using Hyperspectral Imaging.” In: *Journal of Imaging Science and Technology* 65.5, 050406-1–050406–13



*Journal of Imaging Science and Technology*® 65(5): 050406-1–050406-13, 2021.


# Influence of Acquisition Parameters on Pigment Classification using Hyperspectral Imaging

**Dipendra J. Mandal, Sony George, and Marius Pedersen**▲
*Department of Computer Science, Norwegian University of Science and Technology (NTNU), Norway*
*E-mail: dipendra.mandal@ntnu.no*

**Clotilde Boust**
*Centre de Recherche et de Restauration des Musées de France (C2RMF), & CNRS PCMTH PSL*

**Abstract.** *Pigment classification of paintings is considered an important task in the field of cultural heritage. It helps to analyze the object and to know its historical value. This information is also essential for curators and conservators. Hyperspectral imaging technology has been used for pigment characterization for many years and has potential in its scientific analysis. Despite its advantages, there are several challenges linked with hyperspectral image acquisition. The quality of such acquired hyperspectral data can be influenced by different parameters such as focus, signal-to-noise ratio, illumination geometry, etc. Among several, we investigated the effect of four key parameters, namely focus distance, signal-to-noise ratio, integration time, and illumination geometry on pigment classification accuracy for a mockup using hyperspectral imaging in visible and near-infrared regions. The results obtained exemplify that the classification accuracy is influenced by the variation in these parameters. Focus distance and illumination angle have a significant effect on the classification accuracy compared to signal-to-noise ratio and integration time.* © *2021 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.2021.65.5.050406]

## 1. INTRODUCTION

Hyperspectral Imaging (HSI), also called imaging spectroscopy, is a non-invasive imaging technique that generates a spatial map over continuous spectral bands, producing a three-dimensional datacube i.e., two spatial and one spectral dimension. On the basics of data acquisition methods, a spectral data can be created using three general approaches namely, whiskbroom (point scanning), pushbroom (line scanning), and snapshot (single-shot). The line scanning approach is widely adopted because of its higher Signal-to-Noise Ratio (SNR) and flexibility [1]. In this approach, the object is scanned line by line at a time, it uses an array of detectors to scan over a two-dimensional surface using a detector perpendicular to the surface of an object being scanned [2, 3]. HSI technology which was initially developed and used for remote sensing applications [4] has later been used in different application domains such as agriculture [5], medical [6], forensic [7], biomedical engineering [8], Cultural Heritage (CH) [9], etc. Materials with distinct spectra as each element emits a distinctive set of discrete wavelengths according to its atomic and molecular electronic structure [10].

With the development of sophisticated hardware and software, this imaging technology is being used more frequently for the analysis of work of art [11, 12]. Pigment classification of artwork materials such as paintings is of importance for conservators to do a precise analysis of an object and understand its historical value [13, 14]. Despite the significant utilization of HSI in this field [15–18], there are still important challenges in terms of delivering high-quality spectral data. Defining image quality is a complex subject. For three-channel (RGB) imaging, quality criteria are often subjective as it reflects the visual perception of a human observer [19, 20]. However, for HSI it is not only limited to perceptual quality, as it captures data beyond the visible range and is used for a wide range of applications, therefore it is difficult to generalize the definition of quality. Several definitions of spectral quality can be found in the literature and most of them depend upon the application. Fryskowska et al. [21] define quality as the suitability of a specific dataset for a specific purpose. This is more appropriate for spectral imaging in general and thus for pigment classification, the obtained spectral data will be considered to have high quality if the classification accuracy is high. From the perspective of image quality in spectral imaging, most of the research work has been focused on the remote sensing application [22], where the acquisition is made from satellites and aircrafts with significant ground sample distance. The sun is a primary source of illumination; the scattering and absorption of sunlight by different layers of the atmosphere can result in intensity modifications and spectral variations, resulting in degradation of obtained data quality. To deal with such degradation, many algorithms have been developed, for example, correction of atmospheric interference [23]. However, for CH applications, the acquisition is carried out in a close range and the quality of obtained data depends upon parameters of the HSI system such as illumination geometry, the focus of optics, sensor integration time, and SNR [18, 24]. Although instrument calibration is an essential step to obtain valuable and relevant results from HSI, however, more of these parameters are quantified by

▲ IS&T Member.

Received June 4, 2021; accepted for publication Sept. 21, 2021; published online Oct. 14, 2021. Associate Editor: Markku Hauta-Kasari.
1062-3701/2021/65(5)/050406/13/$25.00

device manufacturers and, therefore, are not considered in this paper.

In an artwork, the pigments used are usually mixed, i.e., one or more pigments are mixed with a binder such as oil, egg tempera, gum, etc., and therefore have heterogeneous structures that can have a significant effect on gloss levels. For example, mixture of some pigments, or paintings with varnish layers can cause specular reflection, especially due to the angle of illuminations [25], which will further affect the spectral accuracy and result in incorrect pigment classification. The illumination used is an important factor in imaging. A study done by Toque et al. [26] showed that the spectral reflectance obtained using multispectral imaging of a painting was influenced by the lighting conditions. Intensity, type of illumination, and angle of incidence were key elements that influence the resulting data. In a painting, surfaces are often uneven, therefore acquiring images at an optimal focus distance can be a difficult task. Hence, it is important to explore how this influences the obtained spectral data and classification accuracy.

In CH applications, objects are sensitive to temperature. Any object exposed to an illuminant for a longer time during acquisition can result in a change in the material property due to the heat generated by the illuminant, causing significant damage to the object. One possible way to minimize this effect is by increasing the speed of the acquisition. Higher integration time also increases the noise level in the data. It is very common to use silicon-based detectors, such as a charge-coupled device (CCD) in VNIR HSI systems. Such sensors have lower sensitivity at the two ends of wavelength in the VNIR region, i.e. near 400 and 1000 nanometers (nm), resulting in a noisy spectrum in this region. Thus, it is important to investigate how the reconstructed spectrum of artwork materials differs from their original when there is variation in imaging parameters. Therefore, the objective of this research work is to investigate the effect of imaging parameters such as focus, integration time, SNR, and illumination geometry on the classification accuracy of pigments. The rest of this paper is structured as follows, Section 2 describes the state of the art for image acquisition parameters, and it summaries how these parameters can affect the overall quality when used with the HSI systems for CH applications. Object details, imaging technology, and the experimental framework used are stated in Section 3. Section 4 covers the result with discussions. Finally, Section 5 presents conclusions followed by future work.

## 2. STATE OF THE ART

In a digital imaging system, the acquisition stage can be considered as an essential component. For acquisition of high-quality digital data, several acquisition parameters need to be addressed and controlled. Digital image capture is a function of the light source, reflective surface, distance, the angle between the device, the surface, and the illuminant. The optical resolution, noise, depth of field, integration time, illumination, etc. are some of the important acquisition parameters [27, 28]. These parameters are linked to quality attributes such as sharpness, color, tonality, and resolution, and can influence the overall quality of the captured data. In an imaging device with a low depth of field, the objects at different depths from the camera may appear out of focus if they are away from the focus plane [29, 30]. When capturing CH objects, it should be a sharp focus across the entire object being captured, but depending upon the depth of field and object irregularities, it can result in variation in image sharpness resulting in a blurry image and consequently degrading the quality [31].

Illumination is an important factor that often influences image quality attributes such as color reproduction and texture. During an acquisition, if the object is overexposed, the image will be brighter, and the details of the highlights in the scene will be lost, while on the other hand if the exposure is insufficient, the details in the shadows of the scene will be difficult to distinguish. Loss of image details can reduce the usefulness of the acquired images in CH documentation. Accurate color reproduction is an essential requirement for documentation and study of artworks [32–34] e.g., monitoring the fading phenomena and studying color change due to removal of the varnish layer [35, 36]. Image acquisition parameters mentioned above can be more or less important depending upon the application and objective of imaging. Different imaging technologies can be used for image acquisition, allowing more or less similar acquisition parameters. These imaging technologies can be grouped in multi-band, multispectral, and hyperspectral depending on the number of bands selected over a given spectral interval and on their bandwidths.

Numerous studies have shown the successful use of HSI in the study and analysis of CH artefacts [9, 25, 37, 38]. However, the image acquisition of artworks using HSI has several issues for acquiring high-quality data [39–41] and it involves a number of calibrations and corrections steps to obtain an accurate spectral data [24, 42, 43]. Kubic et al. [25] discussed some problems of HSI acquisition of a painting. Depth of field is also crucial for close-range HSI, particularly for artwork such as paintings that are often warped or have uneven surfaces. Thus, acquiring spectral data at the optimum focus can be challenging. Qureshi et al. [41] discussed few challenges involved in the acquisition and processing of HSI for documents. SNR, integration time, and illumination are the most highlighted imaging parameters that influence the quality of HSI data. Pillay et al. [24] have addressed similar parameters and the usefulness of filters, such as equalization and polarizing filters, in the HSI acquisition workflow that can affect the overall data quality.

To gain a better understanding of spectral imaging devices and analyze how they influence data reliability for different artworks, working group 1 of the EU COST-Action TD1201, Color and Space in Cultural Heritage (COSCH) (website: http://www.cosch.info/) initiated a round-robin test that was carried out by nineteen institutions across Europe for five different types of objects using both multispectral and HSI [17]. It addresses various issues related

to instrumentation, data collection, and post-processing over the accuracy and reliability of data. The resulting data was affected (error in spectral alignment, noise, spatial distortion, etc.) by various aspects, such as device configuration, acquisition environment, and methods of data processing, and this could further have an important effect on pigment classification. MacDonald et al. [44] performed a quality assessment of Russian icon digitization, it was one among five different objects used in the COSCH project. They found that the obtained data was degraded due to specular reflections from both glossy painted and metallic gold areas of the icon's surface indicating the control over the illumination geometry. The imaging system used and workflows employed by the participating institutions varied widely, including camera specifications, illumination, imaging geometry, and file formats. It also highlighted that there is a strong need of guidelines for the spectral imaging workflow.

Generally, halogen lamps are used as an illumination source due to their continuous spectrum of light, from ultraviolet to mid-infrared region, i.e., 350–3400 nm. Halogen lamps emit significant levels of electromagnetic radiation, a lot of energy is converted into heat. Organic materials are more sensitive to heat, moderate heat can change the properties of varnish affecting the glossiness of an object. Artwork exposed under excessive heat for a longer time can have a destructive impact, for instance, melting of the varnish or even the paint layer [45]. This can be minimized by following the guidelines, which suggest the use of a proper illumination level (150–200 lux for oil paintings and 50 lux for manuscripts and other paper-based artworks) and other environmental factors for sensitive CH objects [46, 47]. An illumination source that raises the surface temperature of an object more than four degrees Fahrenheit (257.6 K) in the total acquisition process, is not recommended [48]. Fundamentally light-induced damage is determined by the accumulated total energy incident on material i.e. lux hour rather than the intensity of the incident light. Illumination used in an imaging device setup for an art object can be used either with low-intensity light for a longer time or with high-intensity light for a short time; in both conditions, we may achieve similar SNR. Nevertheless, this reciprocity principle might not always hold for every work of art, for example, some pigments in a painting, and can be independent of time period [49]. However, due to the total energy incident on an object, higher intensity light is preferred [38].

Whetton et al. [50] evaluated the effect of camera height, angle, integration time, and distance between the illuminant and the object on the SNR for wheat plant canopy captured with an HSI system and found these parameters to have a high influence on the spectral quality. A noisy spectrum was obtained when imaged with low integration time and a larger distance between an illuminant and the object. Due to the acquisition setup similarity (i.e., close range), we assume these parameters might also influence the acquisition of CH applications. Likewise, Wang et al. [51] also mentioned focus and integration time as factors influencing spectral image quality. For acquisition of images of fruits using HSI, it was difficult to preserve the focus due to nonuniform fruit size (parameters comparable to CH objects), resulting in either too bright or too dark areas within the fruits making feature extraction a difficult task. To solve this issue, the author recommended a few steps such as changing the orientation of scanning, adding additional lamps, and using a multi-step reflectance target.

Researchers often prepare mockups using specific pigments mixed using binders [52, 53]. These are modern pigments having similar properties to historical pigments from different periods. Generally, shiny materials were used in traditional Asian arts and imaging such objects often causes serious challenges as the intensity of the specular reflection component is usually much higher than that of the diffuse reflection, producing a saturated image. Light scattering is dependent upon the surface properties such as roughness, reflective binders, varnishes, etc., and can modify the spectral reflectance behavior [54, 55], it can also cause specular reflection especially on varnished or glossy paintings. Even in controlled laboratory conditions, non-homogeneously illuminated paintings result in highlights and shadowy areas and degrade the overall quality below a useful level.

In artwork analysis, one of the important tasks is pigment identification [26, 38, 56, 57]. For pigment classification using HSI, the two common approaches are supervised and unsupervised methods. Researchers mostly use supervised classification [58, 59], where they compare the obtained spectrum with a reference spectrum that is mostly created within a Region Of Interest (ROI) and stored as a spectral library, whereas, in an unsupervised method, it looks for spectral clustering of pixels [59, 60]. One of the most commonly applied classification algorithms using data from HSI is the Spectral Angle Mapper (SAM) [61, 62]. This method considers the angle formed between the spectrum of the reference and the test image at each pixel, where smaller angles represent a closer match of the spectrum. Each spectrum is treated as a vector in an $N$-dimensional space where $N$ is equal to the number of spectral bands. Few other algorithms used for supervised classification are spectral correlation mapper [62], maximum likelihood [63], spectral information divergence [64], and spectral gradient mapper [65]. It is essential to assure that the spectral data acquired from the artwork is accurate to achieve the precise classification of the materials/pigments present in the artwork. The HSI acquisition parameters can influence the quality of the spectral data, and the objective of the presented research is to study the link between them.

## 3. MATERIALS AND METHODS

In this section, we describe the test object and the hyperspectral image acquisition laboratory setup, followed by details on the acquisition parameters. The classification model and data post-processing steps are also explained.

### 3.1 *Test Object*

A pigment mockup [53] was used as a test object in this work. The reason to choose this mockup as an object was because

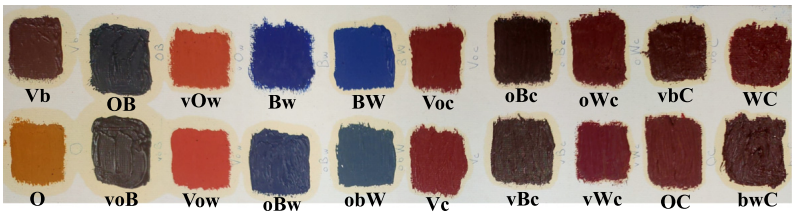



Figure 1. Pigment mockup used as test object. Labels for patches have been added here for description and are not part of the mockup.

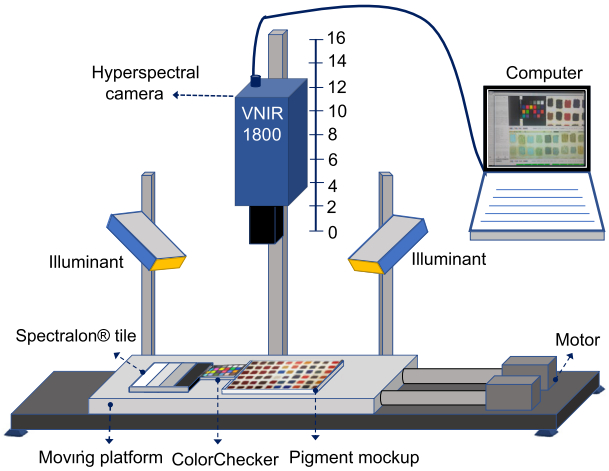



Figure 2. The layout of HSI system used for the experiment. Illumination used was a 150 W halogen-based SmartLite 3900e produced by Illumination Technologies, Inc., guided on the object via optical fiber. Illumination geometry is 45°- 0°- 45°, here 0° implies camera angle with normal.

of its material and physical characteristics considering its usefulness in CH. Powder pigments that are known to have been used in the historic period (14th–18th century) were mixed using linseed oil as a binder and applied over a stretched canvas that were pre-primed using gesso. Patches were made using different concentrations of seven pigments, each weighed on a precision scale. The pigments were Vermilion (V), Ultramarine Blue (B), Viridian Green (G), Naples Yellow (Y), Gold Ochre (O), Kremer White (W), and Novoperm Carmine Red (C). In the remaining part of this paper, we will denote these pigments with their abbreviations. Abbreviations in the capital and small letter will be used to denote the concentration of mixtures. For example, the letter VB denotes that the ratio of mass is 1:1 for pigments B and V. Similarly, Bv is 2:1 and Bvy means 2:1:1. A picture of the pigment mockup is shown in Figure 1.

### 3.2 *Experimental Setup*

Hyperspectral images were obtained in a laboratory environment using the line scanner HySpex VNIR-1800 developed by Norsk Electro Optikk [66], consisting of an actively cooled and stabilized complementary metal-oxide-semiconductor detector. The spectral data obtained covers a spectral range from 400 to 1000 nm with 186 spectral bands having a spectral sampling of 3.26 nm. The scanning speed is automatically synchronized with the integration time which is manually set on the device using the camera interface software HySpex GROUND. In this experiment, a 30 cm cylindrical lens was used that captures 1800 spatial pixels across a line with a field of view of approximately 86 mm.

As shown in Figure 2 the experiment was conducted in a laboratory environment and a translation stage setup was used where the pigment mockup was placed on the moving

platform. The Spectralon® multi-step reference target [67] consisting of four adjacent panels with reflectance values 99, 50, 25, and 12% and a ColorChecker [68] was also kept along with the test target at the same horizontal level in every scan as shown in Figure 3 and were perpendicular to the focal axis of the camera [69]. The reference target is used for computing the normalized reflectance at the pixel level. The objective of using a ColorChecker was to validate the obtained spectral data.

### 3.3 *Methodology*

Setup as shown in Fig. 2, the optimal focus is obtained at a distance of 22 cm from the camera as claimed by the HSI device manufacturer. We will consider this distance as a reference focus point (Gnd_T) throughout this paper. For focus, we choose to change the distance away from the camera with a step size of 2 cm from Gnd_T. Due to the arrangement of the setup, it was convenient to move the camera in the direction as shown in Fig. 2. The Number of scans ($N$) was changed for pushbroom HSI, which is scanning every single line multiple times and taking an average before moving to the following line. This procedure improves the SNR ratio by a factor of $N$. For a work of art reducing the measurement time as much as possible reduces the exposure to the radiation during acquisition of HSI data which further helps in safeguarding the analyzed work. Therefore, for SNR, acquisition with a value of $N$ equal to 1, 2, 4, 6, and 8 was carried out and it was done by giving input directly to the software provided by device manufacture. Orientation including other acquisition parameters was kept as specified in Fig. 2.

Integration time is another important attribute of image acquisition. Acquiring an image at a lower integration time will make the acquisition process faster and lower the exposure of an object to the illumination. There is a trade-off between light intensity and integration time, as it is important to keep the art object less exposed to high light intensities. Therefore, for this part of the experiment, we changed the integration time from the minimum (allowed by the device software i.e., 2150 µs) to a certain higher value (i.e., 12,500 µs) so that its pixels have saturation values between 85% and 10%. The scanning was conducted with SNR equal to 2. In the last part of the experiment, we studied the influence of illumination angle on the acquired spectral data for classification accuracy. The standard configuration for scanning is at 0°, 45° for the camera and illuminant, respectively. We changed the angle of the illuminant to 30° and 60°.

Detectors have low sensitivity at low and high extremes of the spectral range and the illumination intensities near these regions are weak as well, thus resulting in the adding of noise in the spectral data. One possible way is to use an equalization filter. This helps to improve the SNR mainly towards the extremes of wavelength at the same time it also limits the power efficiency of the light source in the central region of the detector and might need a longer integration time. We used an equalization filter on the device and the acquisition of the pigment mockup was carried out at an illumination angle of 45° to observe its effect on the obtained spectral data.

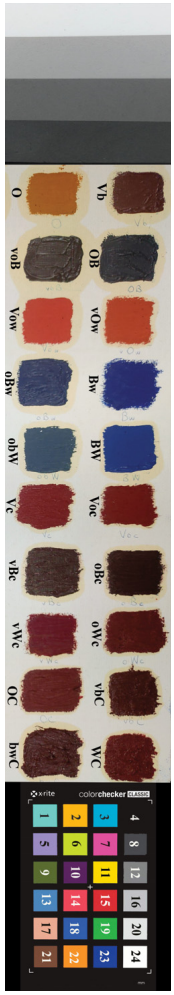



Figure 3. Acquisition arrangement of pigment mockup with the Spectralon® multi-step reference target to the left and ColorChecker on the right. Numbering for ColorChecker is added manually here in this figure for reference.

### 3.4 *Data Processing*

The obtained raw hyperspectral data require post-processing to acquire calibrated normalized reflectance data. Radiometric calibration was carried out where the raw digital number

data from the camera was corrected for non-uniformity and dark offset and then converted to sensor level absolute radiance value using the standalone post-processing software HySpex RAD. Finally, the reflectance factor for the pigment mockup was calculated using the known reflectance value of the reference target. Calculation is shown in Eq. (1), where, $R_{\mathrm{Obj}}(\lambda)$ is the reflectance of an object, $R_{\mathrm{Ref\_}t}(\lambda)$ is the reflectance of reference target, $r_{\mathrm{Obj}}(\lambda)$ and $r_{\mathrm{Ref\_}t}$ are sensor absolute radiance values for the object and reference target, respectively. The reference target surface might have some variation in pixel value, so we averaged the values from 100 pixels for each line scan and calculated the reference target radiance value. Due to the small distance between the sensor and the object, we assumed that the path radiance effect to be negligible. The obtained spectral data was then cropped to exclude the ColorChecker and the reference target. The modification was made using the open-source software Spectralpython [70].

$$R_{\mathrm{Obj}}(\lambda) = R_{\mathrm{Ref\_}t}(\lambda)\frac{r_{\mathrm{Obj}}(\lambda)}{r_{\mathrm{Ref\_}t}(\lambda)} \tag{1}$$

$$\alpha = \cos^{-1}\frac{\sum_{i=1}^{nb} t_i r_i}{\sqrt{\sum_{i=1}^{nb} {t_i}^2}\sqrt{\sum_{i=1}^{nb} {r_i}^2}}. \tag{2}$$

For classification, a supervised approach using the SAM algorithm was applied with a default threshold angle of 0.1 radians. The spectral angle between an image pixel and reference spectrum is given by Eq. (2), where $\alpha$ is the spectral angle in radians, $t_i$ is the image spectrum, $r_i$ is the reference spectrum and $nb$ is the total number of bands. We defined the training region for each of the pigment patches, i.e. an ROI of approximate size equal to that of the patches (25 × 25 mm) was considered, and the regional mean spectrum from these patches were stored and used as the reference spectrum. The classification accuracy was calculated with the statistical parameters, i.e., confusion matrix [71, 72] using the commercial remote sensing software Environment for Visualizing Images (ENVI). The overall methodology is illustrated using a block diagram in Figure 4 and Table I shows the summary of the acquisition parameters.

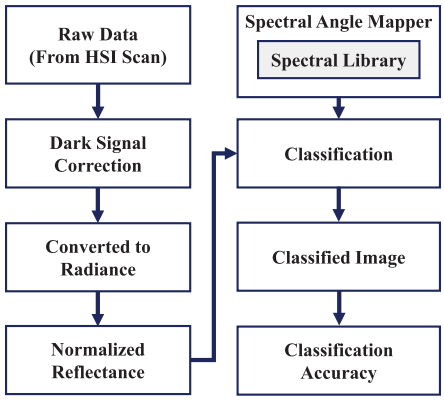



Figure 4. Hyperspectral data processing workflow diagram.

Figure 5(a) shows the pigment patches and Fig. 5(b) illustrates its corresponding image after classification. Different colors in Fig. 5(a) indicate pixels for the particular patch. Accuracy is evaluated as the ratio of classified pixels to the total pixels in a given ROI polygon. As an example, the result of the classification of four patches under optimal acquisition conditions is shown in Table II.

## 4. RESULTS AND DISCUSSION

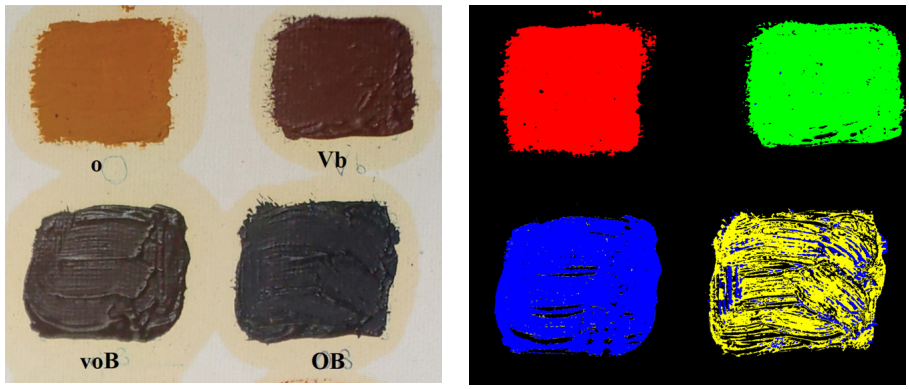
In this section, we will look in detail at the spectrum and classification accuracy obtained for the mockup and ColorChecker by varying the quality attributes, i.e. focus, SNR, integration time, and illumination angle.

### 4.1 *Focus*

Figure 6 shows the spectrum of three different patches for varying focus distance from 0 cm (Gnd_T) to 16 cm away from the initial position of the camera. For the patch O (Fig. 6a) there are slight changes in the magnitude of the spectrum mainly in the range between 600 and 1000 nm. Whereas for the patch OB (Fig. 6b) and patch voB (Fig. 6c) we can see a spectral variation in both visible and near-infrared regions. The spectrum is plotted for a small region within the given patch, i.e., averaging 10 × 10 pixels, each patch is

**Table I.** Acquisition parameters. Variable indicates the different values at which acquisition was done and fixed parameters imply the condition that was constant for each set of experiments; I is illumination measured.

| Acquisition parameters | Variables | Fixed parameters |
|---|---|---|
| Focus distance (F) | {Gnd_T, 2, 4, 6, 8, 10, 12, 14, 16} cm | SNR = 2, IT = 12,500 µs, A = 45°, and I = 3200 lux |
| SNR | {*N* = 1, 2, 4, 6, 8} | F = Gnd_T, IT = 12,500 µs, A = 45°, and I = 3200 lux |
| Integration Time (IT) | {2150, 2500, 5000, 7500, 10,000, 12,500} µs | F = Gnd_T, SNR = 2, A = 45°, and I = 3200 lux |
| Illumination Angle (A) | 30°, 45° and 60°, and I = 2375, 3200 and 4700 lux | F = Gnd_T, SNR = 2, and IT = 12,500 µs |




(a) (b)

Figure 5. (a) Pigment patches, single pigment O; Vb patch is a mixture of two with concentration 2:1; voB mixture of three with concentration 1:1:2; OB contains two pigments with equal concentration. (b) Classified image, where color indicates the class that each pixel has been classified to.

**Table II.** Classification accuracy: total number of pixels classified correctly for each patch within the selected ROI.

| | Patch O Red | Patch Vb Green | Patch voB Blue | Patch OB Yellow |
|---|---|---|---|---|
| Patch O (Red) | **80,850 (99.98%)** | 0 | 0 | 0 |
| Patch Vb (Green) | 0 | **73,223 (99.90%)** | 0 | 0 |
| Patch voB (Blue) | 01 | 0 | **68,337 (93.40%)** | 30,739 (33.50%) |
| Patch OB (yellow) | 0 | 0 | 2159 (2.95%) | **45,812 (49.93%)** |
| Unclassified | 14 (0.02%) | 74 (0.10%) | 2668 (3.65%) | 15,209 (16.57%) |
| Total No. of Pixels | **80,864** | **73,297** | **73,164** | **91,760** |

approximately 500 × 500 pixels. In general, we can observe that there is a change in the magnitude of the spectrum, and the shape of the spectrum is moreover constant.

It was also seen that variation in magnitude of spectrum change with the number of pixels chosen to average, in fact, data plotted from sub-areas in different places within the same patch showed high variations as shown in Figure 7(a). This is mainly because of variation in pigments mixture concentration and nonuniformity in the applied layers. The effect of this variation and nonuniformity is also seen in classification, as shown in Fig. 5(b), not all patches are equally classified and thus have different classification accuracy. More pixels are classified in patches with a single pigment and or homogeneous texture compared to that of having a rough texture. Liang [38] also mentioned that the ratio of pigment concentration to binding medium affects the peak of the spectrum. This argument can be supported by observing the spectra of ColorChecker patches as shown in Fig. 7(b), we can observe that there is a slight variation in magnitude towards the higher wavelength but still the overall shape of the spectrum is similar.

The result for the pigment classification overall accuracy for the given mockup is shown in Figure 8. It is observed that the classification accuracy initially increases as an object gets further away from the camera starting from the optimal focus point and after some points, it starts to decrease. It is because as it moves away from an object, pixels become slightly out of focus and therefore blurred (smoothed), and more adjacent pixels are averaged. As the camera moves further away from the pigment mockup, the camera is outside the optimal focus distance and depth of field, thus photons from the pigment patch area are no longer hitting the same pixels and start to hit adjacent pixels and affect the obtained spectrum.

### 4.2 *SNR*

Spectrum for various SNR levels (frame averaging) for three different patches are shown in Figure 9. The three patches shown are a patch with a single pigment (Fig. 9a), a mixture

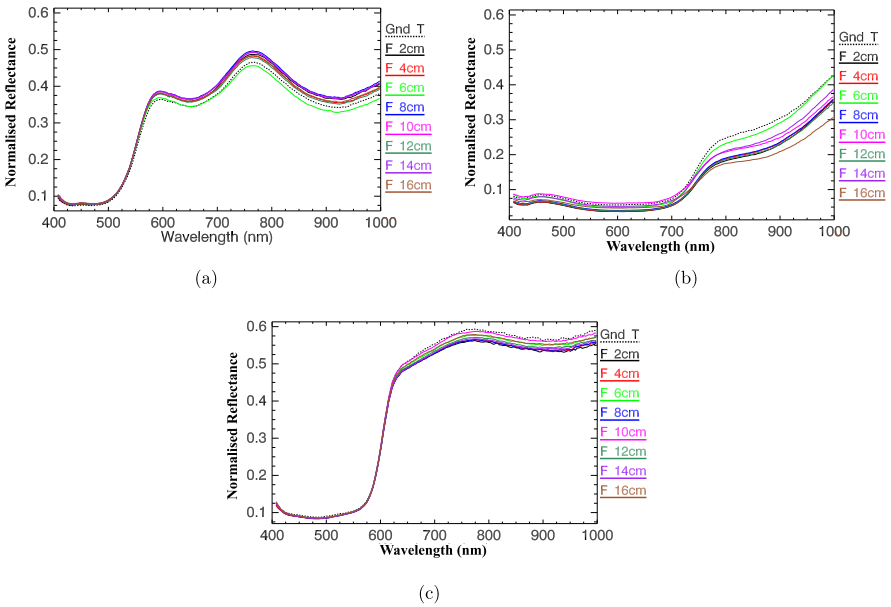



Figure 6. Spectrum for different pigment patch at different focus distance; (a) Patch O, (b) Patch OB, and (c) Patch Vow.

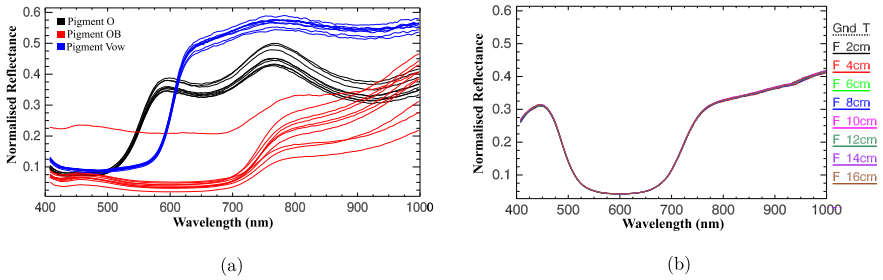



Figure 7. (a) Spectrum for pigment patches obtained at multi-point; for patch O (in black) and patch Vow (in blue), variation is mainly from 550–1000 nm whereas, for patch OB (in red), variation is over the entire wavelength range with slightly higher towards 1000 nm; (b) Spectrum for ColorChecker blue patch (number 23) at the different focus distance.

of two (Fig. 9b), and a mixture of 3 pigments (Fig. 9c). To get a smoother spectral curve, we used a window size of $10 \times 10$ pixels i.e., averaging the spectrum over 10 adjacent pixels. It is observed that the change in the spectrum for different SNR levels was lower for all patches over the entire wavelength range. There is no recommendation or standard practice for considering an exact number of pixels to plot the average spectrum. However, experts recommend focusing on a small section of paintings by using between 6 and 18 pixels. When the size of this window was changed to $1 \times 1$ pixel, we notice a variation in spectrum i.e., noisier which decreased with a higher value of $N$, a result for a patch O is shown in Fig. 9(d). Spectrum is plotted with an offset in normalized reflectance for better visualization.

It can be observed that there is a variation in the spectrum mostly in the regions of 400–500 nm and 850–1000 nm and as SNR increases the spectrum become smoother (less variation). The variation seen is for the reason

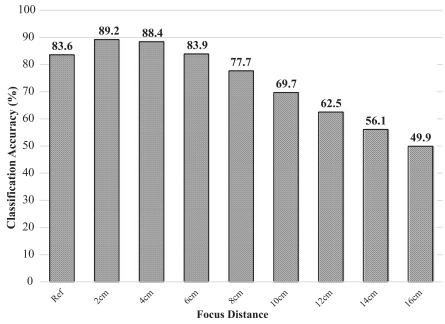



Figure 8. Overall classification accuracy for pigment mockup at the different focus distances.

that exposure in this region requires a longer time compared to mid-region wavelength to overcome the effect of lower quantum energy (sensitivity) of a detector. One possible way to improve this is by using an equalization filter, as it attenuates the light in the mid-region and improves the relative SNR at both ends region of the wavelength range. Figure 10 illustrates the spectrum obtained after using an equalization filter for three patches and the ColorChecker (white patch #24). The spectrum obtained with and without a filter is shown by a solid and dotted line, respectively. It can be seen that, towards both ends of the wavelength range, the spectral variation decreases when the equalization filter is used. The experiment was performed at a distance of 30 cm from the camera to the mockup and this close distance could be one reason for obtaining less noisy data. Classification accuracy for pigment mockup for the different SNRs did not differ much, and the results are shown in Figure 11, there is only a change in obtained value after the second decimal. We also computed the classification accuracy for the ColorChecker, and it was above 99% for all 24 patches, this is because the standard ColorChecker patches have smooth spectral curve characteristics.

### 4.3 *Integration Time*

For different integration times, as shown in Figure 12, the magnitude of the spectrum for all three patches did not show any significant changes. A slight shift can be noticed for patches OB (Fig. 12b) and Vow (Fig. 12c) in the range of 800–1000 nm. It implies similar accuracy in HSI data can be obtained with reduced measurement time, i.e., less exposure of an object to radiation. As shown in Fig. 12(d), the spectrum obtained from the ColorChecker for different integration times were also identical. During the acquisition, the illumination intensity varied such that pixels in the field of view have saturation values between 85% and 10%, and the spectrum was plotted by taking an average of $10 \times 10$ pixels. However, for the non-homogeneous paintings, neighboring pixels could have different characteristics. Thus, obtaining an average over a bigger window size would not be possible, so that the result might be affected. The classification accuracy is shown in Figure 13, which illustrates that for the variation in integration time, the classification accuracy is moreover the same. It can also be observed that for variation in either of parameters SNR or integration time, similar classification accuracy can be obtained with sufficient illumination level.

### 4.4 *Illumination*

To analyze the effect of illumination angle on the spectral data, the acquisition was done with focused illuminants at three different angles. The result obtained is shown in Figure 14. It illustrates that there is a shift in the spectrum in the range of 600–1000 nm for patch O and Vow, but this is not the case for Vb and OB. The reason for this is assumed to be the non-homogeneity in the surface of pigment patches. It can also be noticed that at an angle of 30° and 45° there is very little variation in the spectrum for all four cases whereas, for O and Vow, there is a slightly high shift in magnitude at 60°. Classification accuracy, as shown in Figure 15, is higher at 45° and changes at a different angle of illumination.

### 4.5 *General Observation*

Patches in a pigment mockup are separated from each other, which is not common in real paintings, as elements in the paintings are normally close to each other. Adjacent pixels of different pigments can be misclassified as the spatial resolution changes with focus distance and can change the classification result. A shift in the spectrum does not have any significant effect on the classification accuracy for the attributes SNR and integration time, but this could be important for other applications such as fading or applications that have different concentrations of the same pigment. The influence of SNR and integration time can be more visible if the distance between the camera and object is increased, which could be the case when scanning larger paintings. For larger objects, a rotational stage is used which introduces geometrical errors. An experiment can be conducted in the future to see how this geometrical distortion affects the classification accuracy. Despite having optimal instrumental setup and calibration workflow, pigments surface non-homogeneity in artwork arising from the brushstroke, various thickness layers, compositions in pigments, etc., can affect the obtained data, resulting in misclassification and identification. Further experiments need to be conducted to analyze various factors constituting non-homogeneity on pigment surfaces in works of art, for example, thicknesses, textures, etc. and correlating them with classification accuracy. The mockup in our experiment was unvarnished, usually, paintings are varnished in a real scenario and illumination geometry can cause specular reflection on painting [73]. Berns et al. [74] explained in detail about optics behind varnished paintings, which states that the physical parameters of a varnish affect its optical properties when applied to paintings. Experiments with a new mockup addressing these limitations can be conducted in the future to get a comprehensive result for classification accuracy.

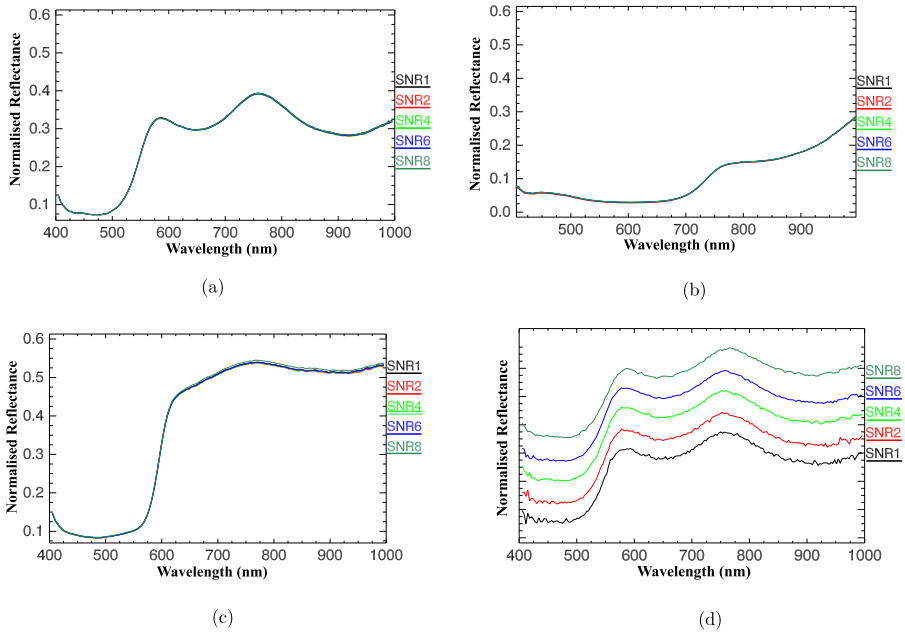



Figure 9. A spectrum of different pigments patch for different SNR; (a) Patch O, (b) Patch OB, (c) Patch Vow, and (d) Spectrum for patch O with offset in the normalized reflectance.

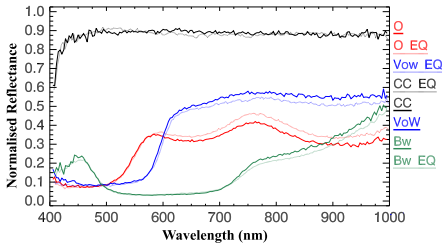



Figure 10. Spectrum for different pigment patches and the ColorChecker: with and without an equalization filter. The letters O, Vow, and Bw represent spectrum for pigment without equalization filter and letter with underscore suffix EQ represents spectrum obtained using the equalization filter.

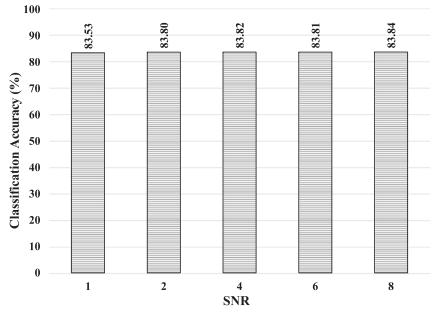



Figure 11. Classification accuracy at different SNR values.

## 5. CONCLUSION

Hyperspectral imaging is being used more frequently in the cultural heritage field to study materials and their distribution. The quality of the acquired hyperspectral data is important to produce accurate and reproducible spectral data for the analysis and documentation of a work of art. It can be influenced by different acquisition parameters and is also dependent upon the attributes linked to specific applications. In CH, pigment classification of artwork materials, such as paintings, is of importance for conservators for precise analysis of objects and their historic value. Therefore, to understand how the acquisition

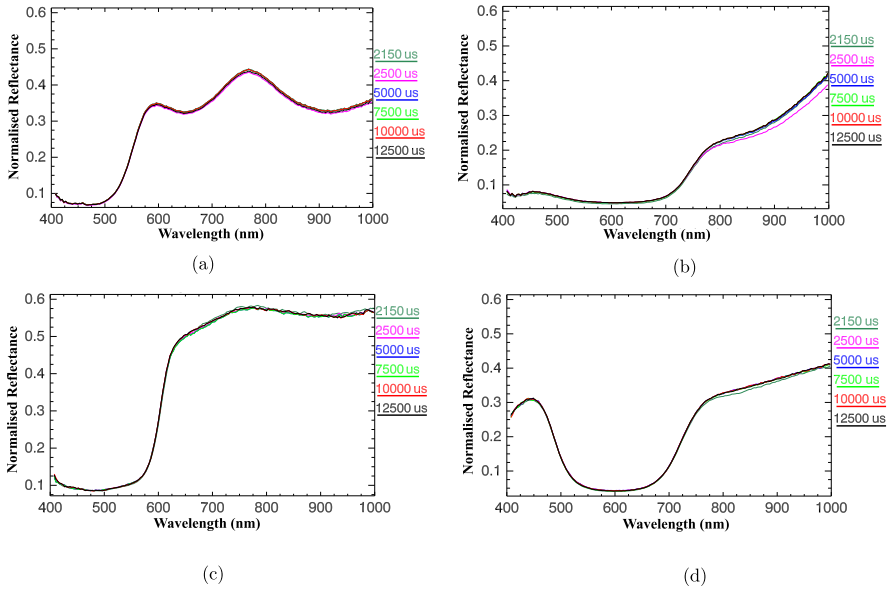



Figure 12. Spectrum for different pigment patches and ColorChecker for different integration time; (a) Patch O, (b) Patch OB, (c) Patch Vow and (d) ColorChecker blue patch (number 23).

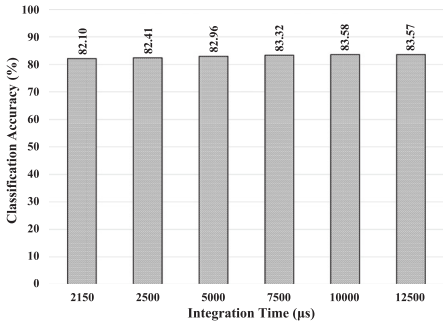



Figure 13. Classification accuracy at different integration time.

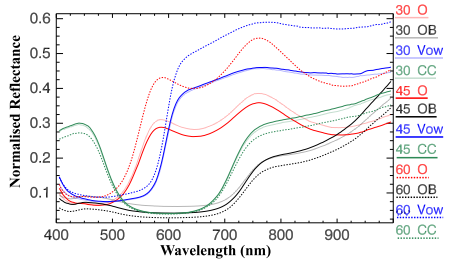



Figure 14. Spectrum for different pigments patches and the ColorChecker at three different illumination angles. The number in the legend represents the angle of illumination.

parameters affect the quality of the obtained spectral data, we investigated the influence of four key parameters, namely, focus distance, signal-to-noise ratio, integration time, and illumination geometry on pigment classification accuracy for a mockup using hyperspectral imaging in visible and near-infrared regions.

We observed that pigment classification accuracy is influenced by a change in focus distance. Moving an object away from the focus plane, pixels appear out of focus resulting in a blurred image. Blurring acts as a low pass filter and smooths edges and consequently increases the classification accuracy, however, after a certain distance, the classification accuracy starts to decrease. SNR and integration time have less effect over classification compared to focus. One possible reason for this might be due to less noise in a close-range laboratory setup. The pigment patches in the mockup have an uneven surface, which results in

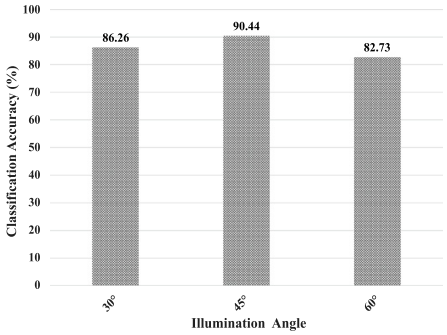



Figure 15. Classification accuracy at a different angle of illumination.

significant variation in the spectrum obtained at different pixels within the same patch. Changing the illumination angle changes the magnitude of the obtained spectrum to some extent and also varies the classification accuracy. An equalization filter can help to reduce the noise in the obtained spectrum especially at two ends of the wavelength range in the VNIR region.

## ACKNOWLEDGMENT

This work is carried out at the Norwegian Colour and Visual Computing Laboratory (Colourlab), within the Department of Computer Science (IDI), as part of the CHANGE (Cultural Heritage Analysis for New Generations) project. And has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 813789.

## REFERENCES

[1] D. Coulter, P. L. Hauff, and W. L. Kerby, "Airborne hyperspectral remote sensing," *Proc. 5th Decennial Int'l. Conf. on Mineral Exploration* (Decennial Mineral Exploration Conferences, Toronto, Ontario, Canada, 2007), pp. 375–378.

[2] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification* (Kluwer Academic/Plenum Publishers, Dordrecht, Netherlands, 2003).

[3] J. M. Amigo and H. Babamoradi, "Saioa Elcoroaristizabal. Hyperspectral image analysis. A tutorial," Anal. Chim. Acta **896**, 34–51 (2015).

[4] F. D. van der Meer, H. M. A. van der Werff, F. J. A. van Ruitenbeek, C. A. Hecker, W. H. Bakker, M. F. Noomen, M. van der Meijde, E. J. M. Carranza, J. B. de Smeth, and T. Woldai, "Multi- and hyperspectral geologic remote sensing: A review," Int. J. Appl. Earth Observat. Geoinform. **14**, 112–128 (2012).

[5] L. M. Dale, A. Thewis, C. Boudry, I. Rotar, P. Dardenne, V. Baeten, and J. A. F. Pierna, "Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review," Appl. Spectrosc. Rev. **48**, 142–159 (2013).

[6] G. Lu and B. Fei, "Medical hyperspectral imaging: A review," J. Biomed. Opt. **19**, 1–24 (2014).

[7] G. Edelman, E. Gaston, T. van Leeuwen, P. J. Cullen, and M. C. G. Aalders, "Hyperspectral imaging for non-contact analysis of forensic traces," Forensic Sci. Int. **223**, 28–39 (2012).

[8] Q. Li, X. He, Y. Wang, H. Liu, D. Xu, and F. Guo, "Review of spectral imaging technology in biomedical engineering: achievements and challenges," J. Biomed. Opt. **18**, 1–29 (2013).

[9] C. Fischer and I. Kakoulli, "Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications," Stud. Conserv. **51**, 3–16 (2006).

[10] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," IEEE Signal Process. Mag. **19**, 12–16 (2002).

[11] F. G. France, "Advanced spectral imaging for noninvasive microanalysis of cultural heritage materials: Review of application to documents in the U.S. library of congress," Appl. Spectrosc. **65**, 565–574 (2011). PMID: 21639977.

[12] C. Cucci, J. K. Delaney, and M. Picollo, "Reflectance hyperspectral imaging for investigation of works of art: old master paintings and illuminated manuscripts," Acc. Chem. Res. **49**, 2070–2079 (2016).

[13] R. Mayer, *The Painter's Craft: An Introduction to Artists' Methods and Materials*, Studio Book (Van Nostrand, New York, NY, 1966).

[14] B. H. Stuart, "Conservation materials," *Analytical Techniques in Materials Conservation* (John Wiley & Sons, Ltd., 2007), Chapter 1, pp. 1–42.

[15] D. Bai, D. W. Messinger, and D. Howell, "Hyperspectral analysis of cultural heritage artifacts: pigment material diversity in the Gough Map of Britain," Opt. Eng. **56**, 1–11 (2017).

[16] M. Picollo, C. Cucci, A. Casini, and L. Stefani, "Hyper-spectral imaging technique in the cultural heritage field: New possible scenarios," Sensors **20**, 2843 (2020).

[17] S. George, J. Y. Hardeberg, J. Linhares, L. Macdonald, C. Montagner, S. Nascimento, M. Picollo, R. Pillay, T. Vitorino, and E. Keats Webb, "A study of spectral imaging acquisition and processing for cultural heritage," *Digital Techniques for Documenting and Preserving Cultural Heritage* (Amsterdam University Press, Amsterdam, 2019), pp. 141–158.

[18] D. Foster and K. Amano, "Hyperspectral imaging in color vision research: Tutorial," J. Opt. Soc. Am. A **36**, 606 (2019).

[19] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process. **13**, 600–612 (2004).

[20] P. G. Engeldrum, "A theory of image quality: The image quality circle," J. Imaging Sci. Technol. **48**, 447–457 (2004).

[21] A. Fryskowska and J. Stachelek, "A no-reference method of geometric content quality analysis of 3D models generated from laser scanning point clouds for hBIM," J. Cultural Heritage **34**, 95–108 (2018). Technoheritage 2017.

[22] E. Christophe, D. Léger, and C. Mailhes, "Quality criteria benchmark for hyperspectral imagery," IEEE Trans. Geosci. Remote Sensing **43**, 2103–2114 (2005).

[23] B.-C. Gao, C. Davis, and A. Goetz, "A review of atmospheric correction techniques for hyperspectral remote sensing of land surfaces and ocean color," *2006 IEEE Int'l. Symposium on Geoscience and Remote Sensing* (IEEE, Piscataway, NJ, 2006), pp. 1979–1981.

[24] R. Pillay, J. Hardeberg, and S. George, "Hyperspectral imaging of art: Acquisition and calibration workflows," J. Am. Inst. Conservat. **58**, 3–15 (2019).

[25] M. Kubik, "Chapter 5 hyperspectral imaging: A new technique for the non-invasive study of artworks," in *Physical Techniques in the Study of Art, Archaeology and Cultural Heritage*, edited by D. Creagh and D. Bradley (Elsevier, Amsterdam, Netherlands, 2007), Vol. 2, pp. 199–259.

[26] J. A. Toque, M. Komori, Y. Murayama, and A. Ide-Ektessabi, "Analytical imaging of traditional japanese paintings using multispectral images," *Int'l. Conf. on Computer Vision, Imaging and Computer Graphics* (Springer, Berlin/Heidelberg, 2009), pp. 119–132.

[27] C. Relf, *Image Acquisition and Processing with LabVIEW*. 07 2003.

[28] M. W. Burke, *Image Acquisition: Handbook of Machine Vision Engineering: Volume 1* (Elsevier, Amsterdam, Netherlands, 1996).

[29] S. Ray, "Chapter 6 – photographic and geometrical optics," in *The Manual of Photography*, edited by E. Allen and S. Triantaphillidou (Focal Press, Oxford, 2011), pp. 103–117.

[30] E. Webb, S. Robson, and R. Evans, "Quantifying depth of field and sharpness for image-based 3D reconstruction of heritage objects," ISPRS – Int'l. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XLIII-B2-2020**, 911–918 (2020).

[31] K. Martinez and A. Hamber, "Towards a colorimetric digital image archive for the visual arts," Proc. SPIE **1073**, 114–121 (1989).

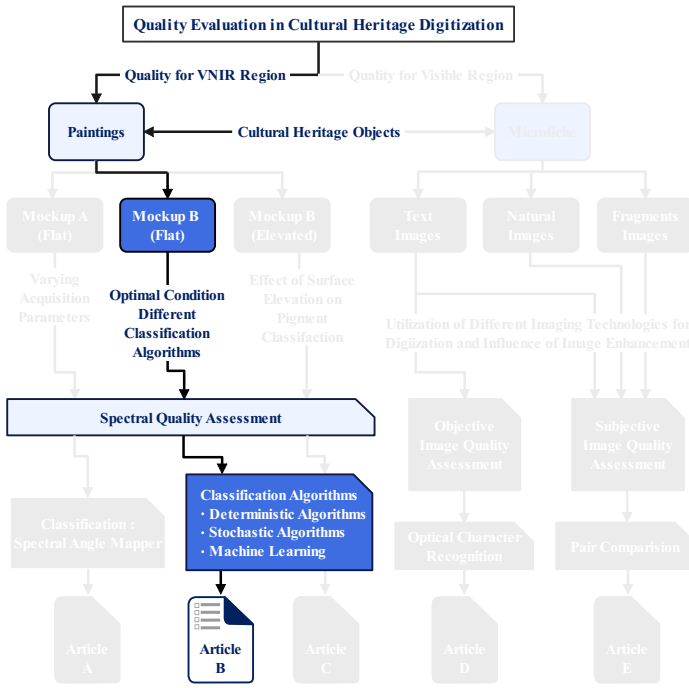
[32] S. Lorusso, A. Natali, and C. Matteucci, "Colorimetry applied to the field of cultural heritage: examples of study cases," Conservation Science in Cultural Heritage 1974-4951 **7**, 187–208 (2007).

[33] P. Korytkowski and A. Olejnik-Krugły, "Precise capture of colors in cultural heritage digitization," Color Res. Appl. **42**, 333–336 (2017).
[34] E. Franceschi, P. Letardi, and G. Luciano, "Colour measurements on patinas and coating system for outdoor bronze monuments," J. Cultural Heritage **7**, 166–170 (2006).
[35] M. Bacci, "Optical spectroscopy and colorimetry," *Proc. Int'l. School of Physics Enrico Fermi* (IOS Press, Ohmsha, 1999, 2004), Vol. 154, pp. 1–16.
[36] R. Fontana, A. D. Fovo, J. Striova, L. Pezzati, E. Pampaloni, M. Raffaelli, and M. Barucci, "Application of non-invasive optical monitoring methodologies to follow and record painting cleaning processes," Appl. Phys. A **121**, 957–966 (2015).
[37] S. Legrand, F. Vanmeert, G. Van der Snickt, M. Alfeld, W. De Nolf, J. Dik, and K. Janssens, "Examination of historical paintings by state-of-the-art hyperspectral imaging methods: From scanning infra-red spectroscopy to computed x-ray laminography," Heritage Sci. **2**, 13 (2014).
[38] H. Liang, "Advances in multispectral and hyperspectral imaging for archaeology and art conservation," Appl. Phys. A: Mater. Sci. Process. **106**, 309–323 (2012).
[39] T. Vitorino, A. Casini, C. Cucci, A. Gebejesje, J. Hiltunen, M. Hauta-Kasari, M. Picollo, and L. Stefani, "Accuracy in colour reproduction: Using a colorchecker chart to assess the usefulness and comparability of data acquired with two hyper-spectral systems," *Computational Color Imaging* (Springer International Publishing, Cham, 2015), pp. 225–235.
[40] F. Daniel, A. Mounier, J. Pérez-Arantegui, C. Pardos, N. Prieto-Taboada, S. Fdez-Ortiz de Vallejuelo, and K. Castro, "Hyperspectral imaging applied to the analysis of goya paintings in the museum of Zaragoza (Spain)," Microchem. J. **126**, 113–120 (2016).
[41] R. Qureshi, M. Uzair, K. Khurshid, and H. Yan, "Hyperspectral document image processing: Applications, challenges and future prospects," Pattern Recognit. **90**, 12–22 (2019).
[42] P. Geladi, J. Burger, and T. Lestander, "Hyperspectral imaging: calibration problems and solutions," Chemo Metr. Intell. Lab. Syst. **72**, 209–217 (2004). Advances in Chromatography and Electrophoresis – Conferentia Chemometrica 2003, Budapest.
[43] J. Qin, "Chapter 5 – hyperspectral imaging instruments," in *Hyperspectral Imaging for Food Quality Analysis and Control*, edited by D.-W. Sun (Academic Press, San Diego, 2010), pp. 129–172.
[44] L. W. MacDonald, T. Vitorino, M. Picollo, R. Pillay, M. Obarzanowski, J. Sobczyk, S. Nascimento, and J. Linhares, "Assessment of multispectral and hyperspectral imaging systems for digitisation of a Russian icon," Heritage Sci. **5**, 1 (2017).
[45] T. R. Peery and D. W. Messinger, "Spatial resolution as a trade-space for low-light imaging of sensitive cultural heritage documents," J. Cultural Heritage **45**, 81–90 (2020).
[46] British Standards Institution. *PAS 198: 2012: Specifications for Managing Environmental Conditions for Cultural Collections* (BSI Standards, London, 2012).
[47] Illuminating Engineering Society and Illuminating Engineering Society of North America. *Recommended Practice for Museum Lighting: ANSI/IES RP-30-17*. ANSI/IES (Illuminating Engineering Society of North America, New York, 2017).
[48] Still Image Working Group. Technical guidelines for digitizing cultural heritage materials. Technical Report, Federal Agencies Digitization Guidelines Initiative (September 2016).
[49] D. Saunders and J. Kirby, "Light-induced damage: investigating the reciprocity principle," *11th Triennial Meeting, Edinburgh, Scotland, 1–6 September, 1996: Preprints* (ICOM Committee for Conservation, Paris, France, 1996), pp. 87–90.
[50] R. Whetton, T. Waine, and A. Mouazen, "Optimising configuration of a hyperspectral imager for on-line field measurement of wheat canopy," Biosyst. Eng. **155**, 84–95 (2017).
[51] N.-N. Wang, D.-W. Sun, Y.-C. Yang, H. Pu, and Z. Zhu, "Recent advances in the application of hyperspectral imaging for evaluating fruit quality," Food Anal. Methods **9**, 178–191 (2015).
[52] A. Hayem-Ghez, E. Ravaud, C. Boust, G. Bastian, M. Menu, and N. Brodie-Linder, "Characterizing pigments with hyperspectral imaging variable false-color composites," Appl. Phys. A **121**, 939–947 (2015).
[53] F. Grillini, J. Thomas, and S. George, "Linear, subtractive and logarithmic optical mixing models in oil painting," *Colour Visual Comput. Symp.* (CEUR-WS.org, Gjøvik, Norway, 2020), Vol. 2688.
[54] I. Sandu, M. Sá, and M. Pereira, "Ancient "gilded" art objects from european cultural heritage: A review on different scales of characterization," Surf. Interface Anal. **43**, 1134–1151 (2011).
[55] T. Cavaleri, A. Giovagnoli, and M. Nervo, "Pigments and mixtures identification by visible reflectance spectroscopy," Proc. Chem. **8**, 45–54 (2013). YOuth in the COnservation of CUltural Heritage, YOCOCU 2012.
[56] R. J. H. Clark, "Pigment identification on medieval manuscripts by Raman microscopy," J. Molecular Struct. **347**, 417–427 (1995). Molecular Spectroscopy and Molecular Structure, 1994.
[57] J. Delaney, E. Walmsley, B. Berrie, and C. Fletcher, "Multispectral imaging of paintings in the infrared to detect and map blue pigments," *Scientific Examination of Art: Modern Techniques in Conservation and Analysis* (The National Academies Press, Washington, DC, 2005), pp. 120–136.
[58] H. Deborah, S. George, and J. Y. Hardeberg, "Pigment mapping of the scream (1893) based on hyperspectral imaging," *Int'l. Conf. on Image and Signal Processing* (Springer, Cham, 2014), Vol. 8509, pp. 247–256.
[59] C. Balas, G. Epitropou, A. Tsapras, and N. Hadjinicolaou, "Hyperspectral imaging and spectral classification for pigment identification and mapping in paintings by El Greco and his workshop," Multimedia Tools Appl. **77**, 9737–9751 (2018).
[60] L. Tan and M.-l. Hou, "A study on the application of SAM classification algorithm in seal of calligraphy and painting based on hyperspectral technology," *2016 4th Int'l. Workshop on Earth Observation and Remote Sensing Applications (EORSA)* (IEEE, Piscataway, NJ, 2016), pp. 415–418.
[61] C. Cucci, A. Casini, L. Stefani, M. Picollo, and J. Jussila, "Bridging research with innovative products: a compact hyperspectral camera for investigating artworks: a feasibility study," Proc. SPIE **10331**, 1033106 (2017).
[62] O. A. De Carvalho and P. R. Meneses, "Spectral correlation mapper (SCM): an improvement on the spectral angle mapper (SAM)," *Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18* (JPL Publication, Pasadena, CA, 2000), Vol. 9.
[63] H. Z. Mohd Shafri, A. Suhaili, and S. Mansor, "The performance of maximum likelihood, spectral angle mapper, neural network and decision tree classifiers in hyperspectral image analysis," J. Comput. Sci. **3**, 419–423 (2007).
[64] C.-I. Chang, "An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis," IEEE Trans. Inform. Theory **46**, 1927–1932 (2000).
[65] E. Angelopoulou, S. Lee, and R. Bajcsy, "Spectral gradient: a material descriptor invariant to geometry and incident illumination," *Proc. Seventh IEEE Int'l. Conf. on Computer Vision* (IEEE, Piscataway, NJ, 1999), Vol. 2, pp. 861–867.
[66] Norsk Elektro Optikk. http://www.hyspex.no/. Accessed: 20 December 2020.
[67] Spectralon® Multi-step Targets. https://www.labspherestore.com/product-p/aa-006xx-000.htm. Accessed: 11 September 2020.
[68] X-Rite ColorChecker Classic. https://www.xrite.com/. Accessed: 11 September 2020.
[69] D. Saunders and J. Cupitt, "Image processing at the national gallery: The vasari project," National Gallery Technical Bulletin **14**, 72–85 (1993).
[70] Welcome to Spectral Python (SPy). http://www.spectralpython.net.
[71] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," Inform. Process. Manage. **45**, 427–437 (2009).
[72] J. Lever, M. Krzywinski, and N. Altman, "Points of significance: Classification evaluation," Nat. Methods **13**, 603–604 (2016).
[73] D. C. Day, Technical report: Evaluation of optical flare and its effects on spectral estimation accuracy. pp. 1–19, 2003. Accessed: 20 December 2020.
[74] R. S. Berns and E. R. de la Rie, "Exploring the optical properties of picture varnishes using imaging techniques," Stud. Conserv. **48**, 73–82 (2003).

# Article B

## An Experiment–based Comparative Analysis of Pigment Classification Algorithms using Hyperspectral Imaging

D. J. Mandal, M. Pedersen, S. George, H. Deborah, and C. Boust (2023). "An Experiment–based Comparative Analysis of Pigment Classification Algorithms using Hyperspectral Imaging." In: *Journal of Imaging Science and Technology* 67.3, pp. 030403-1–030403-18. DOI: 10.2352/J.ImagingSci.Technol.2023.67.3.030403



*Journal of Imaging Science and Technology*® 67(3): 030403-1–030403-18, 2023.


# An Experiment-based Comparative Analysis of Pigment Classification Algorithms using Hyperspectral Imaging

**Dipendra J. Mandal▲, Marius Pedersen▲, Sony George▲, and Hilda Deborah▲**
*Department of Computer Science, Norwegian University of Science and Technology (NTNU), Norway*
*E-mail: dipendra.mandal@ntnu.no*

**Clotilde Boust**
*Center for Research and Restoration of Museums of France (C2RMF), France*

**Abstract.** *Hyperspectral imaging techniques are widely used in cultural heritage for documentation and material analysis. Pigment classification of an artwork is an essential task. Several algorithms have been used for hyperspectral data classification, and the effectiveness of each algorithm depends on the application domain. However, very few have been applied for pigment classification tasks in the cultural heritage domain. Most of these algorithms work effectively for spectral shape differences and might not perform well for spectra with differences in magnitude or for spectra that are nearly similar in shape but might belong to two different pigments. In this work, we evaluate the performance of different supervised-based algorithms and few machine learning models for the pigment classification of a mockup using hyperspectral imaging. The result obtained shows the importance of choosing appropriate algorithms for pigment classification.* © *2023 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.2023.67.3.030403]

## 1. INTRODUCTION

Hyperspectral Imaging (HSI) technology, initially developed and used for remote sensing applications, is also being used more frequently in the Cultural Heritage (CH) domain for analyzing artwork and has provided great potential in its scientific analysis. In CH, proper pigment classification of artwork materials such as paintings is of essential importance for conservators to precisely analyze an object and its historical value. Generally, reflection, transmission, and absorption of electromagnetic energy by a given material produce a unique spectrum at a given wavelength. The shape of the spectrum is distinctive because every material has a different chemical composition and an inherent physical structure [1]. For pigment classification using HSI, supervised classification algorithms are mainly used; they compare the spectrum within a region of interest with spectral library spectra with a specific tolerance [2, 3].

Many supervised-based classification algorithms exist for HSI, mostly in remote sensing applications, for example, mineral identification [4, 5]. However, few of these algorithms are being adopted directly or with some modification in other application domains such as medical imaging [6, 7], food and agriculture [8–10], forensics [11]. Moreover, to the best of our knowledge, only a few have been implemented in the CH domain, especially for pigment classification of artwork such as paintings. HSI acquisition for CH are usually performed under controlled laboratory conditions, where the distance between the camera and the object is relatively small and one has control over illumination types and geometry. In contrast, for remote sensing, HSI data are collected using natural illumination with a more considerable distance between the camera and target, causing temporal illumination variations and atmospheric effects. Due to these differences between two application domains, various classification algorithms adopted in remote sensing cannot be directly adapted or might not work effectively for CH applications. For example, an algorithm insensitive to intensity variation can perform well in remote sensing. However, it might not perform with the same accuracy for CH objects because magnitude measures are essential in CH. Faded or aged pigments [12], pure pigments mixed with different binding mediums [13], mixed pigments (e.g., pigments mixed in different weight percentages of lead white [14]), etc. can have variations in magnitude, which is essential to determine for both diagnostic and conservative purposes. Very few of these algorithms have been used for pigment identification of artwork using HSI, and therefore it is necessary to explore and evaluate them. Furthermore, many materials associated with CH lack pure end members, particularly when they undergo weathering [15], aging [16–18], or restoration processes over time [19]. Therefore, accurately determining the composition of a specific material or differentiating it from other materials within an image can pose challenges, making the task of identifying and mapping materials in HSI more challenging.

Deep learning has recently provided new possibilities by solving more complex questions in many applications [20, 21]. In CH, spectra of the pigments get affected with different types of medium used as binders; spectra might look identical, i.e. might have a small shift in peak or small change in magnitude [13], and under such conditions, most of the supervised algorithms do not perform well for classification. However, distinguishing such conditions might be important

▲ IS&T Members.

Received Jan. 3, 2023; accepted for publication Apr. 5, 2023; published online May 12, 2023. Associate Editor: Rita Hofmann-Sievert.
1062-3701/2023/67(3)/030403/18/$25.00

for art historians and conservators to select the proper conservation methods. Also, in the case of fading, there might only be a minor change in the magnitude of a spectrum. In medical imaging, Zhi et al. [22] used a Support Vector Machine (SVM) for tongue diagnosis using HSI, where spectra obtained from the surface of the tongue under different conditions have changed mainly in magnitude. Devassy et al. [9] used a One-dimensional Convolutional Neural Network (1D-CNN) to classify strawberries and found that the result was better than supervised algorithms. To the best of our knowledge, using deep learning-based models for pigment classification of artwork is not a common practice and, therefore, it will be worthwhile to explore their potential.

This paper presents the comparative experimental analysis of various supervised algorithms and machine learning models for pigment classification on a mockup using HSI in the Visible Near-Infrared (VNIR) region. The algorithms used are the Spectral Angle Mapper (SAM), Spectral Correlation Mapper (SCM), Spectral Information Divergence (SID), Spectral Similarity Scale (SSS), and the hybrid combinations of SID–SAM and SID–SCM. We also used the Jeffries-Matusita (JM) distance function combined with SAM (JM-SAM). Likewise, few of the machine learning models used are SVM, Fully Connected Neural Network (FC-NN), and 1D-CNN. The rest of this paper is structured as follows. Section 2 provides an overview of data processing techniques and algorithms, followed by details about the algorithms used in Section 3. Object details, imaging technology, and the experimental framework used are given in Section 4. Section 5 covers the results with a discussion. Finally, Section 6 presents our conclusions, followed by future work.

## 2. OVERVIEW OF ALGORITHMS AND PROCESSING TECHNIQUES

Generally, a spectral matching technique is employed for pigment classification, i.e., finding a spectral similarity between two spectra at any given pixel in an image. The best fit indicates the most significant possibility of being reference material for a given pixel. The distinction between different algorithms used for classification is their ability to consider shapes and magnitude differences between two spectra. This section provides an overview of the classification algorithms employed in various application domains with HSI.

Shivakumar et al. [23] compared the performance of SAM and SCM for classifying nine different classes in remote sensing applications using HSI. There was spectral overlapping between the datasets for some of the classes, and they identified that SCM was more efficient compared to SAM for the classes with a highly similar spectrum. Similarly, SCM was compared with SAM for mineral analysis [24] and it was found that SCM algorithm delivered better results due to its wide variation of data from −1 to 1. Qin et al. [25] used SID methods to identify lesions in citrus using HSI. Devassy et al. [26] explored the performance of five different algorithms, namely SAM, SCM, ED, SID, and Binary Encoding (BE), for the task of ink classification using HSI. The overall accuracy (average of all inks used) for the SAM algorithm was high compared to all other methods used. None of the methods worked effectively to classify between inks that had nearly similar spectral signatures with only change in magnitude.

For a given two vectors (spectra), Change Vector Analysis (CVA) computes the change in spectral vectors and compares their magnitude with the specified threshold value [27]. It was originally designed for only two spectral dimensions (2 spectral bands), however, using the directional cosine approach, it can be extended to a $N$-dimensional space [28] and is computed using Eq. (1).

$$\alpha_i = \cos^{-1}\left(\frac{t_i - r_i}{\sqrt{\sum_{i=1}^{nb}(t_i - r_i)^2}}\right), \tag{1}$$

where $t_i$ and $r_i$ are the tests and reference image, and $nb$ is the total number of bands with $i = 1, 2, \ldots, nb$. In this method, we will obtain the number of angles $\alpha_i$ equal to the number of bands, which makes the computation complex, details on this explanation and its drawbacks are explained in Ref. [29]. Osmar et al. [29] in their study of change detection methods in a tropical environment using HSI, proposed a new approach to calculate the spectral direction of change using the SAM and the SCM method, and for magnitude, they computed the Mahalanobis distance and the Euclidean distance. The best result was obtained using SAM for similarity and ED for magnitude.

Many hybrid approaches to compute the classification of HSI data have shown improved results in many applications. Using a hybrid approach of SAM and SID was found to produce better results than using them alone [30]. Naresh et al. [31] computed the hybrid of SCM and SID (SID-SCM) for the classification of vigna species and compared their result with the hybrid method of SAM and SID. They performed an experiment for various spectral regions and found that for region 400–700 nm, results are better. Zhang et al. [32] used the hybrid approach by combining Minimum Noise Fraction (MNF) and SAM methods to identify defective tomatoes.

Li et al. [33] proposed a new method called Extended Spectral Angle Mapper (ESAM) for detecting disease in citrus plants for multi- and hyperspectral datasets. The result was compared with supervised methods, Mahalanobis distance, and unsupervised method; k-means and ESAM were found to have better accuracy (86%) than the other two methods (around 64%). Jeffries-Matusita (JM) [34] are mainly used for the separability criterion and optimal band selection, so only the most distinct bands are selected for the data classification task [35, 36]. The JM method is a pairwise distance measure that can be applied mostly to two class cases. Authors have proposed many extensions of JM [37] for use in multiclass classification. The most common is to take the average JM distance computed for all pairs of classes. Deborah et al. [38] evaluated the performance of four

different distance functions named Root Mean Square Error (RMSE), Goodness-of-Fit Coefficient (GFC) [39], Jeffrey divergence, and Levenshtein distance on both synthetic and real hyperspectral datasets to find a suitable distance measure for spectral image processing. They found that for the magnitude change, only RMSE followed by Jeffrey divergence performed in the desired way.

Deborah et al. [40] compared different distance functions for pigment classification tasks on HSI datasets with the presence of spectral noise and variations. Intending to identify the appropriate methods based on suitable selection criteria, they found the Euclidean distance of a Cumulative Spectrum (ECS) to be the most suitable distance function for spectral data. However, in their study, evaluation of these distance functions on artificially simulated spectra and some real spectra from pigment patches from the Kremer pigment chart [41], these charts are screen printed and usually the pigments are in a water-based binder, which might not be the exact representation of the real spectra obtained from an artwork. Bhattacharyya Distance (BD) measures the separability between two classes and has been used in remote sensing applications frequently [42, 43]. BD was used to select the number of bands required for efficient classification, and then SAM and SVM were used for identification of stress symptoms in plants [44].

In recent years, machine learning-based classification methods have been popular and extensively used in many different applications. SVM is one of the machine learning approaches used for classification tasks and has shown efficient results, especially when the training data size is relatively small [5, 45]. Deep learning-based CNN models can learn spectral features more effectively using deeper layers and in many cases, such methods can give us higher classification accuracy than traditional algorithms. Pouyet et al. [46] used the Deep Neural Network (DNN) and compared the result with SAM for pigment identification and mapping using HSI in the SWIR region and found that the DNN model produced better results than SAM. Devassy et al. [9], in their study of strawberry classification based on sugar content, found that algorithms SID and SAM, which rely on the spectrum's geometry, did not perform well, as the two reference spectrum were nearly identical in shape and a small difference in the magnitude of the NIR region of the spectrum. They also showed that 1D-CNN based classification gives better accuracy (96%) compared to SAM (60%) and SID (58%). Table I summarizes the list of algorithms used for HSI data processing, their area of study, and details of the classification/network parameters.

## 3. CLASSIFICATION ALGORITHMS

In this section, we describe the algorithms used in our experiment.

### 3.1 *Euclidean Distance (ED)*

Classification can be computed by calculating the minimum distance between the spectrum to be classified and the reference spectrum of the class. For a given $n$-dimensional image spectrum $t_i$ and a reference spectrum $r_i$, the ED between them is defined using Eq. (2), where $nb$ is the number of spectral bands. ED is proportional to the magnitude of the squared subtractive difference vector, but not its shape [47].

$$\mathrm{ED} = \sqrt{\sum_{i=1}^{nb}(t_i - r_i)^2} \tag{2}$$

### 3.2 *Spectral Angle Mapper (SAM)*

SAM is one of the most popular spectral classification methods used in CH applications due to its easy and rapid approach to mapping spectral similarity. SAM, developed by Boardman [48], measures the spectral similarity between any two spectra (test and reference). Arccosine angles between the two spectra are calculated by treating them as $N$-dimensional vectors in space, where $N$ is equal to the number of spectral bands. The angle between two spectra is calculated using Eq. (3), where $\alpha$ is the spectral angle in radians, $t_i$ is the image spectrum, $r_i$ is the reference spectrum, and $nb$ is the total number of bands. A smaller angle indicates a more decisive match between the spectra. Kruse et al. [48] describe a simplified representation of the spectral angle mapper algorithm using a two-dimensional scatter plot for two band image data. Since the SAM algorithm measures an angle between two vectors and the angle does not change with the length of the vectors, i.e., insensitive to the gain. Therefore, this algorithm does not consider magnitude shifts in the spectrum (see Osmar et al. [24]).

$$\alpha = \cos^{-1}\left(\frac{\sum_{i=1}^{nb} t_i r_i}{\sqrt{\sum_{i=1}^{nb} {t_i}^2}\sqrt{\sum_{i=1}^{nb} {r_i}^2}}\right) \tag{3}$$

### 3.3 *Spectral Correlation Mapper (SCM)*

SCM calculates the Pearson correlation coefficient between two spectra. It standardizes the data, centralizing itself in the mean of the test and reference spectra. By applying arccosine, it can be expressed in angles. This algorithm excludes negative correlation and retains shading effect minimization characteristics similar to SAM, resulting in better classification results [24, 29]. SCM can be computed using Eq. (4), where $\alpha$ is the arccosine of the spectral correlation measure in radians, $t_i$ and $\bar{t}_i$ are the image spectrum and its sample mean, similarly $r_i$ and $\bar{r}_i$ are the reference spectrum and its sample mean, and $nb$ is the total number of bands.

$$\alpha = \cos^{-1}\left(\frac{\sum_{i=1}^{nb}(t_i - \bar{t}_i)(r_i - \bar{r}_i)}{\sqrt{\sum_{i=1}^{nb}(t_i - \bar{t}_i)^2}\sum_{i=1}^{nb}(r_i - \bar{r}_i)^2}\right). \tag{4}$$

### 3.4 *Spectral Information Divergence (SID)*

SID measures spectral similarity between the spectrum of test and reference data for each pixel based on the concept of divergence, i.e. measuring probabilistic discrepancy between

**Table I.** Summary of algorithms used for HSI datasets with its applications and model parameters; Th: threshold value, BS: batch size, LR: learning rate, DR: dropout rate, ReLu: rectified linear unit, HL: hidden layer, CL: convolutional layer, FCL: fully connected layer, KS: kernel size.

| Algorithms | Application | Wavelength | Parameters |
|---|---|---|---|
| ED | Ink classification [26] | 400–1000 nm | — |
| SAM | Pigment classification [49] | 400–1000 nm | Th:0.1 |
| | Mineral classification [50] | 380–2500 nm | — |
| | Ink classification [26] | 400–1000 nm | — |
| | Minerals and land classification [51] | — | — |
| SCM | Ink classification [26] | 400–1000 nm | — |
| | Pigment identification [3] | 370–1100 nm | — |
| | Pigment mapping [52] | 400–2500 nm | Th:0.1 |
| SID | Mineral classification [50] | 380–2500 nm | — |
| | Ink classification [26] | 400–1000 nm | — |
| | Minerals and land classification [51] | — | — |
| | Crops classification [53] | 200–2400 nm | — |
| SSS | Crops classification [54] | — | — |
| SIDSAM | Crop classification [30] | 400–2500 nm | — |
| | Mineral classification [50] | 380–2500 nm | — |
| | Dye and pigment based Inkjet prints [55] | 400–1000 nm | — |
| SIDSCM | Plant classification [31] | 350–2500 nm | — |
| | Mineral classification [50] | 380–2500 nm | — |
| JMSAM | Landcover classification [56] | — | — |
| | Mineral classification [50] | 380–2500 nm | — |
| | Ink classification [26] | 400–1000 nm | — |
| | Dye and pigment based Inkjet prints [55] | 400–2500 nm | — |
| SVM | Tongue diagnosis [22] | 400–1000 nm | — |
| | Crops classification [57] | — | Polynomial Kernel |
| FC-NN | Aerial images classification [58] | — | BS:500, LR: 0.05, DR:0.25, ReLU |
| | Pigment classification [46] | 1000–2500 nm | HL: 4, LR : 0.001, Adam, ReLU/Sigmoid |
| 1D-CNN | Soil texture classification [59] | 400–1000 nm | CL:4, FCL:2, Softmax |
| | Classification of strawberry [9] | 380–2500 nm | Filters: 8, HL: 2, BS:32, KS: 3 |

them. The probability distribution of the test and reference spectra is expressed as Eq. (5) and Eq. (6), respectively [60].

$$p_i = \frac{t_i}{\sum_{i=1}^{nb} t_i} \tag{5}$$

$$q_i = \frac{r_i}{\sum_{i=1}^{nb} r_i}, \tag{6}$$

where, $t_i$ is the image spectrum, $r_i$ is the reference spectrum, and $nb$ is the total number of bands. Using these two probability distributions, SID can be calculated with Eq. (7).

$$\mathrm{SID} = \sum_{i=1}^{nb} p_i \log\left(\frac{p_i}{q_i}\right) + \sum_{i=1}^{nb} q_i \log\left(\frac{q_i}{p_i}\right) \tag{7}$$

### 3.5 *Spectral Similarity Scale (SSS)*

SSS evaluates the shape and magnitude difference between two spectra. Granahan et al. [54, 61] used SSS to analyze hyperspectral atmospheric correction techniques. This

algorithm uses the Euclidean distance metric for magnitude, and correlation for comparing the shape of the spectra. This method combines the calculations of both, giving each an equal weighting [62]. SSS has a scale ranging from a minimum of zero and maximum of the square root of two; smaller the value, the higher the similarity between the spectrum i.e. if two spectrum are collinear then its SSS value will be equal to zero. SSS can be computed using Eq. (8).

$$\mathrm{SSS} = \sqrt{(d_e)^2 + (\hat{r})^2} \tag{8}$$

Here, $d_e$ is the Euclidean distance between two spectra and is computed using Eq. (9) and its value ranges from 0 to 1 due to the factor $1/nb$.

$$d_e = \sqrt{\frac{1}{nb}\sum_{i=1}^{nb}(t_i - r_i)^2} \tag{9}$$

Equation (10) computes the value for $\hat{r}$, where $r$ is the correlation coefficient between the two spectra and is computed using Eq. (11).

$$\hat{r} = (1 - r^2) \tag{10}$$

$$r^2 = \left( \frac{\sum_{i=1}^{nb}(t_i - \bar{t}_i)(r_i - \bar{r}_i)}{\sqrt{\sum_{i=1}^{nb}(t_i - \bar{t}_i)^2 \sum_{i=1}^{nb}(r_i - \bar{r}_i)^2}} \right)^2 \tag{11}$$

### 3.6 *SID-SAM*

As the name suggests, SIDSAM is computed by multiplying SID by taking the tangent of SAM or with the sine function of SAM, i.e., by computing the perpendicular distance between two vectors (test and reference). Both of these measures produce similar results [30]. This hybrid computation makes two similar spectra even more comparable and two dissimilar spectra more distinctive, thus significantly improving the spectral discriminability. SIDSAM can be computed as either of the Eqs. (12) or (13), where SID and SAM can be computed using Eqs. (7) and (3) respectively.

$$\mathrm{SID} - \mathrm{SAM} = \mathrm{SID} * \tan(\mathrm{SAM}) \tag{12}$$

$$\mathrm{SID} - \mathrm{SAM} = \mathrm{SID} * \sin(\mathrm{SAM}) \tag{13}$$

### 3.7 *SID-SCM*

Similar to SID-SAM, we also tested the hybrid combination of SIDSCM, computed by multiplying SID by either taking a tangent of SCM or with the sine function of SCM [31]. SID-SCM can be computed as either of Eqs. (14) or (15), where SID and SCM can be computed using Eqs. (7) and (4) respectively.

$$\mathrm{SID} - \mathrm{SCM} = \mathrm{SID} * \tan(\mathrm{SCM}) \tag{14}$$

$$\mathrm{SID} - \mathrm{SCM} = \mathrm{SID} * \sin(\mathrm{SCM}) \tag{15}$$

### 3.8 *Jeffries-Matusita Spectral Angle Mapper (JM-SAM)*

Similarly to SID-SAM, JM-SAM is also a hybrid similarity measure algorithm in which the spectral capabilities of both algorithms are orthogonally projected by using either a tangent or a sine function [56]. A smaller JM-SAM value indicates a strong match between the reference and test spectra. It can be computed using either Eqs. (16) or (17).

$$\mathrm{JM} - \mathrm{SAM} = \mathrm{JMD} * \tan(\mathrm{SAM}) \tag{16}$$

$$\mathrm{JM} - \mathrm{SAM} = \mathrm{JMD} * \sin(\mathrm{SAM}) \tag{17}$$

Here, Jeffries-Matusita distance (JMD) is one of the spectral separability measures commonly used in remote sensing applications and can be computed using Eq. (18), where B is the Bhattacharyya distance and is computed using Eq. (19) and SAM is computed using Eq. (3).

$$\mathrm{JMD} = 2(1 - e^{-B}) \tag{18}$$

$$B = \frac{1}{8}(\mu_t - \mu_r)^T \left[\frac{\sigma_t + \sigma_r}{2}\right]^{-1} (\mu_t - \mu_r) + \frac{1}{2}\ln\left[\frac{\left|\frac{\sigma_t + \sigma_r}{2}\right|}{\sqrt{|\sigma_t||\sigma_r|}}\right] \tag{19}$$

Here, $\mu_t$ and $\mu_r$ are the mean of the test and reference spectra, respectively; $\sigma_t$ and $\sigma_r$ are the covariance of the test and reference spectra, respectively.

### 3.9 *Support Vector Machine (SVM)*

SVM is a supervised classification algorithm used in machine learning and has been used successfully for HSI classification tasks [63–65]. These are usually used to separate two or more data classes using a hyperplane. Objects to be classified are represented as a vector in an $n$-dimensional space. Then SVM method draws a hyperplane so that all points of one class are on one side of this hyperplane and points of the other class are on the other side. Of course, there could be multiple such hyperplanes. SVM tries to find the one that best separates these classes by computing the maximum distance between the data points of these classes closest to the hyperplane, also called support vectors. This method is similar to the Neural Network, but instead of computing the weight and bases of each point, SVM adjusts these parameters by computing it only on the support vectors and determining the decision boundaries for classification.

### 3.10 *Fully Connected Neural Network (FC-NN)*

In the FC-NN architecture, all the nodes in one layer are connected to the nodes in the next layer. The data are inputted into the first layer of the neural network, where individual neurons pass the data to a second layer. The second layer of neurons does its task, and so on, until the final layer. Each neuron assigns a weight to its input. Once all the input weights flow out of the neuron, they are summed, and biases are added, which help offset the output. These parameters are tuned by optimization during training, that is, compute the error of classification, also called loss, and then tune the weights and biases over many iterations to minimize this loss. The goal of neural networks is to adjust their weights and biases so that they can produce the desired output when applied to new unseen data. One of the common problems when training the network is

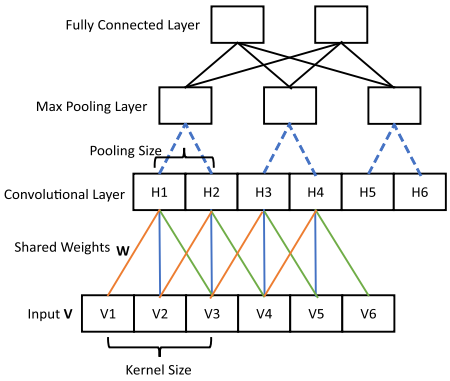



Figure 1. The architecture of a typical CNN consisting of a convolutional layer, a max pooling layer, and a fully connected layer.

overfitting (also called generalization error) of the dataset, i.e., Instead of learning, it memorizes the data. To avoid it, one needs to use regularization, i.e., early stopping with dropout layers and changing the network structure and parameters (weight constraint) [66]. A dropout function added to the network helps to disable the neurons randomly. This forces the network to learn how to make accurate predictions with only randomly left neurons, helping the network to prevent overfitting. For further details, see [67, 68].

### 3.11 *One-dimensional Convolutional Neural Network (1D-CNN)*

CNN is one of the most popular neural networks used for various computer vision and machine learning tasks [69–71]. CNN architecture is built using three main layers: convolutional layer, pooling layer, and fully connected layer. As the name suggests, the convolutional layer performs the linear operation between matrices, that is, convolution between the input neurons and kernel, generating an output activation map. For 1D-CNN, only 1D convolution is performed, that is, scalar multiplications and additions. In this layer, the number of weights is equal to the size of the kernel and does not depend on the input neuron, as in FC-NN. The feature map generated from this layer is passed through pooling a layer which helps to reduce the dimension of the feature map while maintaining the most important information. This helps to introduce translation invariance and reduces overfitting. A fully connected layer takes the output of the pooling layers, flattens them, and turns them into one long vector that can be an input for the next stage, where it applies weights to predict the correct label, and finally outputs the probabilities for each class using the activation function. Figure 1 shows the architecture of a general CNN [72].

## 4. MATERIALS AND METHODS

In this section, we describe the mockup and the HSI acquisition laboratory setup, details on the data post-processing steps, and classification algorithms.

### 4.1 *Test Object*

As shown in Figure 2, a pigment mockup was prepared and used in a laboratory environment. We used pigment tubes composed of high-stability pigments and oil, purchased from Zecchi [73]. The pigments were selected on the basis of the popularity in CH research articles, their spectrum characteristics, and in consultation with experts. Veridian (V), Cerulean Blue (CB), Green Earth (GE), Yellow Ochre Light (YOL), Burnt Umber (BU), Ultramarine Blue Deep (UBD), Lead White Hue (LWH), Genuine Vermilion (GV), Cobalt Blue Deep (CBD), and Ivory Black (IB) are the pigments that are being used in the mockup. The linen canvas used was primed using three layers of white gesso.

### 4.2 *Experimental Setup*

Hyperspectral data were obtained in a laboratory environment using the HySpex line scanner VNIR-1800 from Norsk Electro Optikk [74]. The datacube obtained covers a spectral

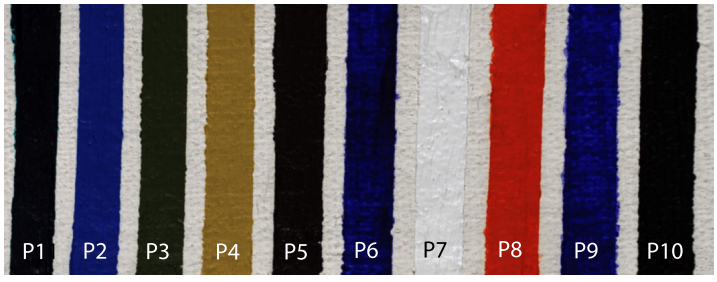



Figure 2. Pigment mockup; P1: Veridian, P2: Cerulean Blue, P3: Green Earth, P4: Yellow Ochre Light, P5: Burnt Umber, P6: Ultramarine Blue Deep, P7: Lead White Hue, P8: Genuine Vermilion, P9: Cobalt Blue Deep and P10: Ivory Black.

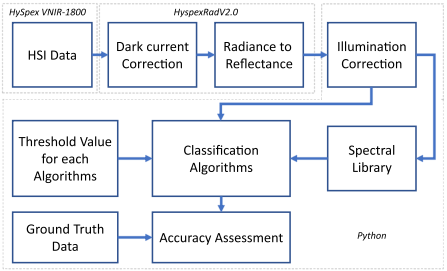



Figure 3. Workflow diagram for data processing.

range from 400 to 1000 nm with 186 spectral bands having a spectral resolution of 3.26 nm. In this experiment, a close-range 30 cm lens was used; it captures 1800 spatial pixels across a linear field of view of approximately 86 mm. A translation stage setup was used where the pigment mockup was kept lying on a horizontal surface. The standard multistep reference target from Spectralon [75] consisting of four shades of 99, 50, 25, and 12% reflectance values was kept along with the mockup during acquisition. This reference target with a known reflectance factors is used for computing the normalized reflectance at the pixel level.

### 4.3 *Data Processing*

The obtained raw hyperspectral data was post-processed for radiometric calibration using the HySpex RAD software, which removes electronics noise, i.e., dark current, and converts the raw images to the sensor absolute radiance values. Illumination correction, i.e., spatial variability in illumination, was performed with the help of the standard reference target. Further data processing steps are different for supervised and ML-based classification and are explained in the following sections.

#### 4.3.1 *Data Processing for Supervised Classification*

To build a spectral library, a region of interest of approximate size equal to that of the patches (10 × 10 mm) was considered, and the mean spectra from these regions were saved in the library. To evaluate the performance of classification, a confusion matrix was computed. The overall methodology is illustrated using a block diagram in Figure 3. All data processing steps were computed using the open-source software Spectralpython [76].

Selecting the appropriate threshold value for classification algorithms is critical as it may vary depending upon the application. For example, Li et al. [33] pointed that the region for selecting the threshold value for SAM to be 0.1 for citrus disease detection analysis because, during the preliminary testing, they found that at a value of 0.15, many false positives result. A similar empirical approach has been followed by Júnior et al. [29], and Fung et al. [77]. Thus we also computed the optimal threshold for each of these algorithms through empirical observation. First, we selected a small segment of the HSI dataset of a mockup, as shown in Figure 4. Next, the reference spectrum was extracted from a mockup's flat region by taking an average of 11 × 11 pixels. Finally, we computed the classification task for all algorithms with different threshold values and evaluated their accuracy using the confusion matrix.

In CH applications such as pigment classification for a painting, misclassification, i.e., the pigment being classified as the wrong pigment, is even more crucial than a pigment being unclassified. Hence, there should be the minimum error for any given classification algorithm. Therefore, we considered the classification accuracy for pigment classified as correct pigment (P_P), misclassification (MC_), pigment classified as unknown (P_UN_), unknown classified as a pigment (UN_P_), and unknown classified as unknown (UN_UN_). Figure 5 shows the graph for these parameters over accuracy for the SID algorithm, and we can observe that for threshold values between 0.01 and 0.03, the accuracy for pigment classified as pigment and unknown classified as unknown is high. Also, for misclassification value in the range of 0.1–0.3, pigments that are classified as unknown is minimum, and unknown classified as unknown is relatively high and constant. A similar conclusion can be drawn by visualizing the classification result shown in Figure 6. An optimal threshold value used for different algorithms in our experiment is mentioned in Table II and graph for each of the algorithms is attached in Appendix B.

#### 4.3.2 *Data Processing for ML Classification*

The obtained normalized reflectance HSI data needs to process before it is fed to the model; data was labeled for different classes using the label encoder. For our dataset, we used one hot encoder, meaning for each class, one value is hot (i.e., the value of 1), and the rest are cold (i.e., the value of 0). We divided the dataset into training and testing. With an 80-20 split, data was further normalized. We then build and implement the model; first training dataset is used to train the model, neural network weights and biases of neurons are updated with each epoch till we got considerably minimum MSE and higher accuracy. Finally, the test dataset is used to

Figure 4. A snippet of a mockup with ten pigments and substrate; Colors are approximated as RGB rendering using spectral python for bands 75, 46, and 19 of HSI datasets.

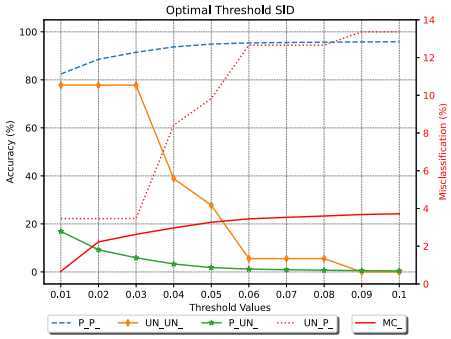



Figure 5. Graph for accuracy of five parameters used to determine the optimal threshold value for the SID algorithm.

**Table II.** The selected threshold value for eight different classification algorithms.

| Algorithms | Threshold value |
|---|---|
| ED | 0.9 |
| SAM | 0.1 |
| SCM | 0.8 |
| SID | 0.03 |
| SSS | 1.1 |
| SID-SAM | 0.003 |
| SID-SCM | 0.005 |
| JM- SAM | 0.09 |

validate the model. A block diagram in Figure 7 illustrates an overall workflow. Training spectra of 10 pigments and a substrate, plotted over a spatial region of approx. 100 × 100 pixels with 186 spectral bands is attached in Appendix A.

SVM model was implemented in Python using the Sklearn library. Among the differences, we tuned our model for three key hyperparameters, namely kernel types, regularization, and gamma, using the Python library called GridSearchCV. This function cross-validates the model to avoid overfitting using *k*-fold cross-validation and then computes a grid to evaluate the performance of each combination of given hyperparameters. Table III shows the details of hyperparameters.

For FC-NN, we build a sequential model with three dense layers, the first layer with 32 nodes and hyperbolic tangent (tanh) as activation function followed by batch normalization. The second layer has 16 nodes tanh activation function followed by batch normalization and dropout, and the third layer has 11 nodes and a softmax activation function. The activation function introduces the non-linearity into the networks so that the networks can learn the relationship between the input and output. Tangent hyperbolic is a non-linear function with an s-shaped graph

**Table III.** SVM key hyperparameters, the range used for tuning, and the optimum value selected for classification; RBF: Gaussian Kernel Radial Basis Function.

| Hyperparameter | Range used | Optimum value selected |
|---|---|---|
| Kernel | "Polynomial", "RBF", "Sigmoid", "Linear" | RBF |
| Regularization | 0.1, 1, 10, 100, 1000 | 100 |
| Gamma | 1, 0.1, 0.01, 0.001 | 1 |
| *k*-fold | 5 | 5 |

with output ranges from −1 to 1. One reason for using the tanh function is that it is zero-centred, which makes the optimization icon process much more manageable. The softmax activation function converts a value vector to a probability distribution and is used in the output layer of multiclass classification. For details on the activation function, please refer [78]. For multiclass classification, the categorical cross-entropy loss function is usually used, and optimization algorithms, which are used to update weights and biases; we used adaptive moment estimation (Adam), as it is the best among the adaptive optimizers in most of the cases [79, 80]. The network architecture used for our experiment is shown in Figure 8. The model was implemented in Python using Keras, a neural network application programming interface.

The proposed 1D-CNN model was tuned for hyperparameters using KerasTuner [81]. We tuned the model for the number of convolutional layers, their filter size, dropout, dense layer filer size, learning rate and epoch. Figure 9 illustrates the block diagram of the tuned model with its hyperparameter used. We used Adam as an optimizer with a learning rate of 0.001 and categorical cross-entropy as the loss function.

## 5. RESULT AND DISCUSSION

This section will look in detail at the classified image, the accuracy obtained for each pigment, and the overall accuracy of the algorithms used. Figure 10 shows the classification accuracy of each pigment for the different algorithms. The classification result for each of these algorithms is attached in Appendix D. We can observe that the average accuracy (average of 10 pigments) is high for all three machine learning algorithms. Of these three, FC-NN has the highest accuracy, followed by 1D-CNN and SVM. For the eight supervised algorithms used, SCM and SAM have high accuracy, followed by SID, SID-SAM, SID-SCM, and SSS. ED and JM-SAM have the lowest classification accuracy.

Apart from machine learning algorithms, the other eight algorithms used have difficulty classifying pigment 6 (P6) and pigment 9 (P9). We can see in Figure 11 that spectra for both these pigments are similar and have little difference in magnitude. This is a common issue with supervision-based classification algorithms [9, 23]. In distance-based algorithms, ED, SSS, and JM-SAM, the classification accuracy for similar spectra (P6 & P9) are the lowest. We also observed that the classification accuracy

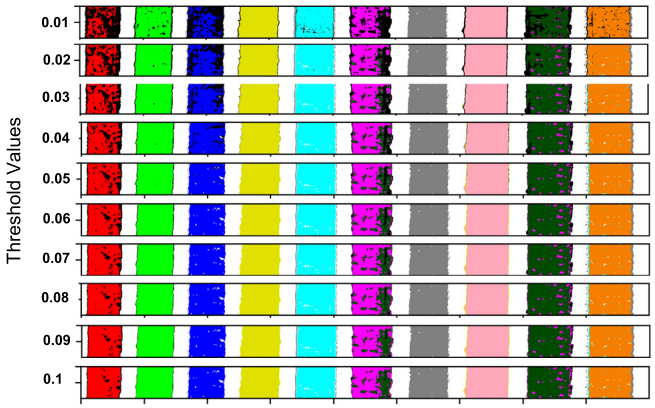



Figure 6. Classification result for ten pigments patches obtained using SID algorithm for a different set of threshold values.

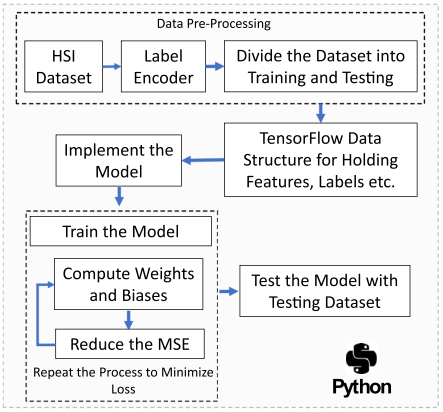



Figure 7. Workflow diagram for ML data processing.

is low for these distance-based algorithms, particularly for pigment 7 (White Hue), which has a spectrum similar to the substrate (S), since it is misclassified as substrate, as shown in the confusion matrix in Figure 12. Pigment 10, as shown in Fig. 11, has a reflectance value below 0.05 for almost the entire wavelength region (450–1000 nm), and it seems that the low magnitude value has an influence on the classification accuracy for supervised-based algorithms. Spectra for all pigments and substrates used are provided in Appendix C.

Classification accuracy for algorithm SID and its hybrid combinations (SID-SAM and SID-SCM) are lower for pigments P1 and P3. Figure 13 shows the classification result for pigments P1, P2, and P3 for SID, SID-SAM, and SID-SCM. Black color represents the unclassified pixels, and we can observe that all three algorithms have similar areas that have not been classified for P1 and P3. From the confusion matrix shown in Figure 14, we can see that for P1 and P3, the unclassified (UC) percentage is the second highest value in all three algorithms.

Figure 15 shows the spectra for reference, classified pixels, and unclassified pixels for pigments P1, P2, and P3. It can be observed that there is a difference in spectra in the range of 800–1000 nm. The solid red line represents the reference spectrum, whereas red dash lines are spectra for classified pixels, and solid green lines are for unclassified pixels for P1. Similarly, the solid blue line is a reference spectrum for P3, and solid orange and solid black lines are spectra for classified and unclassified pixels, respectively. We also plotted the range for P2, which is mostly classified. Dashed blue line is a reference spectrum for P2, and solid grey lines are spectra for classified pixels.

The SID algorithm uses a divergence measure to match the reference and target pixels; the smaller the divergence value, the more likely the pixels are similar. We have used a threshold of 0.03, meaning that pixels with a value less than 0.03 will only be classified, and a value greater than the threshold will not be classified. We computed the divergence value for a spectrum of classified and unclassified pixels with a reference spectrum for P1, P2, and P3. The spectra used in the calculation are shown in Figure 16. The computed divergence is shown in Table IV. We can see that spectra that are not classified in the case of P1 and P3 have divergence values greater than a threshold. We can change this value to get more pixels classified, but this will result in higher misclassification and increase the unknown classified as a pigment, as shown in Fig. 5.

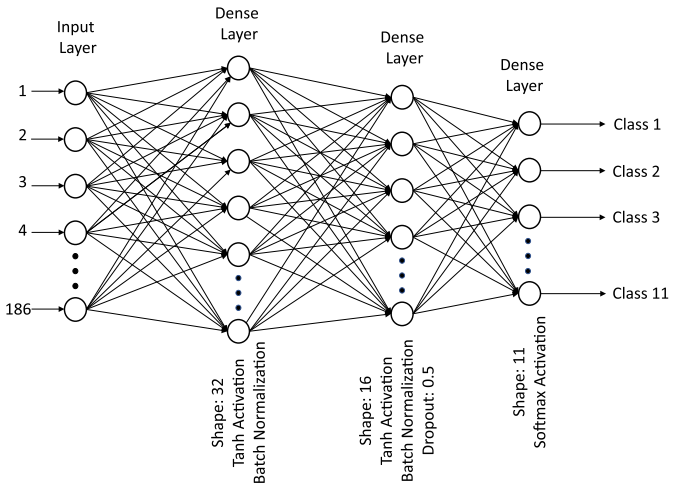



Figure 8. The architecture of the FC-NN classifier used in our experiment.

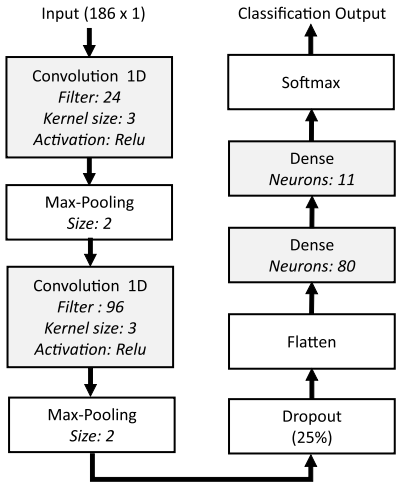



Figure 9. The architecture of tuned 1D-CNN model.

## 6. GENERAL DISCUSSION

Experimental results show that ML algorithms outperform the supervision-based algorithms used. The limitation of supervision-based algorithms used is that they cannot perform well if pigments have nearly identical spectra (P6 and P9) and also if the magnitude of the spectrum is very low (P10, reflectance factor below 0.05). We found for nearly identical spectrum, SCM is a better measure than the SAM, and this could be because SCM considers value from −1 to 1 whereas the cosine of SAM only varies from 0 to 1. Apart from pigments P1 and P3, we found that the SID's hybrid approach with SAM and SCM has almost similar results for our dataset. Due to the threshold value selected for classification, the accuracy for P1 and P3 is lower than for other pigments, i.e., in SID for P1 and P3 threshold value should be greater than 0.3 as mentioned in Table IV. The classification accuracy of algorithms based on spectral distance, such as ED, SSS, and JM-SAM was the lowest. This could be because these algorithms misclassified in-between white pigment (P7) and substrate (S), which is not the case for other supervised algorithms.

**Table IV.** SID value computed between a spectrum of reference pixels with that of classified and unclassified pixels for P1, P2, and P3. Remark indicates that either obtained SID value is smaller or greater than a used threshold value of 0.03.

| Spectra | SID value | Remark |
|---|---|---|
| P1 Ref. & P1 C | 0.005 | <0.03 |
| P1 Ref. & P1 UC | 0.052 | >0.03 |
| P2 Ref. & P2 C | 0.003 | <0.03 |
| P3 Ref. & P3 C | 0.014 | <0.03 |
| P3 Ref. & P3 UC | 0.034 | >0.03 |

ML-based algorithms need to be trained for which we need a large amount of data. Classification result depends upon how well the model is trained, i.e., how large the training datasets are so that model can learn enough distinct features. For ML-based algorithms to perform well and avoid

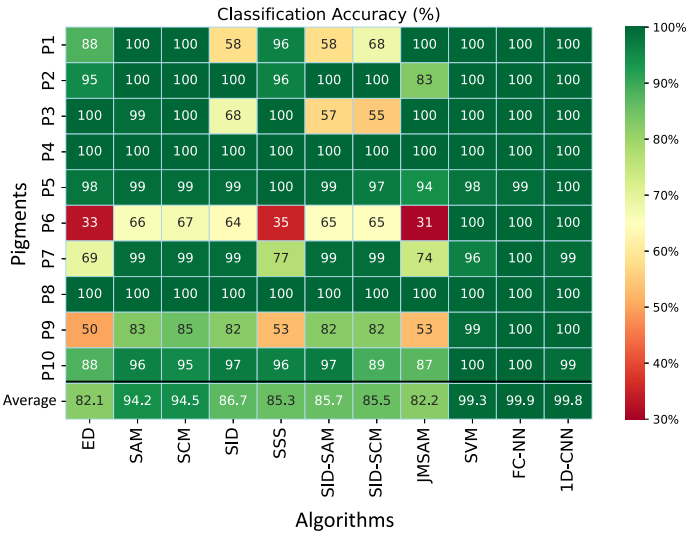



Figure 10. Classification accuracy for each pigment for all 11 algorithms used; average represents the accuracy for an average of 10 pigments for a given algorithm.

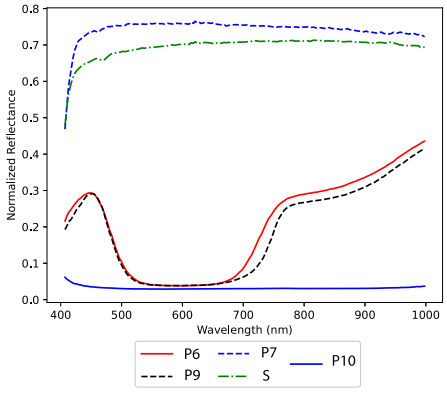



Figure 11. Normalized reflectance spectra for pigment, used as a reference for supervised classification; P6, P7, P9, P10, and S represent pigments 6, 7, 9, 10, and substrate, respectively.

overfitting of a model, it needs to be tuned for the appropriate value of different hyperparameters, which will take a long computing time. This adds to the cost of computational time and complexity for ML-based algorithms. On the other hand, supervision-based algorithms do not require such a training set and are simple and easy to compute. Therefore, for the pigments with less complex spectra (i.e., having less identical spectrum), supervision-based algorithms such as SCM and SAM might be a good fit for the classification task.

## 7. CONCLUSION

HSI is a non-invasive imaging technique used for the documentation and analysis of artwork for various tasks, such as pigment classification. It is essential as it assists conservators and curators in precisely analyzing an object and its historical value. In this paper, we evaluated the spectral processing algorithms for pigment classification of a mockup using HSI. We analyzed eight spectral image classification algorithms, i.e., ED, SAM, SCM, SID, SSS, SID-SAM, SID-SCM, JM-SAM, and three machine learning-based algorithms, SVM, FC-NN, 1D-CNN for its classification accuracy. In general, machine learning algorithms outperformed the others. Supervision-based algorithms work well for the pigments if their spectra are very distinct in shape from each other. Still, these algorithms have poor performance for pigments having a similar spectrum (nearly identical) or spectrum with just a change in magnitude. However, machine learning-based algorithms can overcome this limitation by extracting the features from each training sample and thus perform better for pigment classification. During our experiment, we trained the network for ten pigments. However, extending the model's scope to include a more extensive range of pigments would be beneficial. Additionally, exploring diverse scenarios, such as mixed and aged pigments, would be beneficial; therefore, one can conduct more comprehensive research in the future. By doing so, we can refine the supervised algorithms and machine

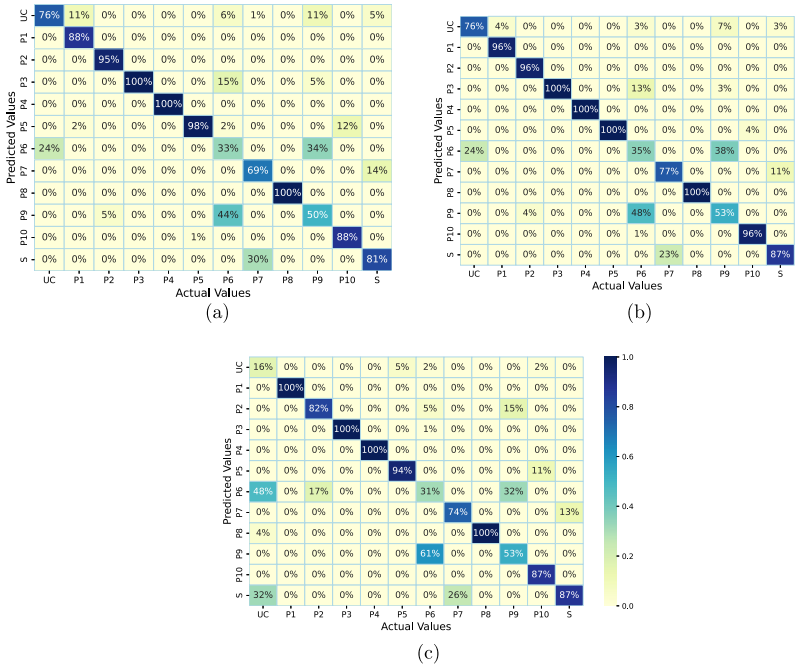



Figure 12. Confusion matrix of (a) ED, (b) SSS, and (c) JM-SAM.

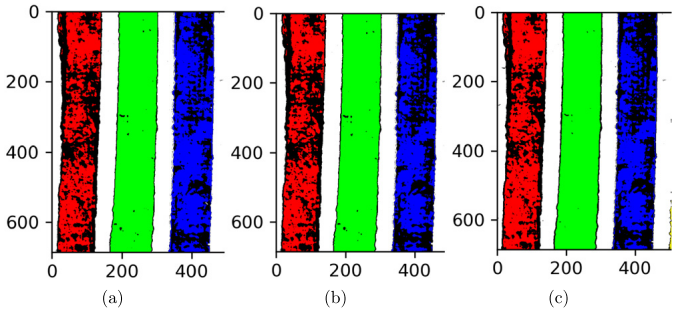



Figure 13. Classification results for pigment P1 (in red), P2 (in green), and P3 (in blue). (a), (b), and (c) are obtained using algorithms SID, SID-SAM, and SID-SCM, respectively.

learning models mentioned earlier to be more applicable to real-world cases in cultural heritage.

**ACKNOWLEDGMENT**

This work is carried out at the Norwegian Colour and Visual Computing Laboratory (Colourlab), within the Department of Computer Science (IDI), as part of the CHANGE (Cultural Heritage Analysis for New Generations) project. And has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 813789.

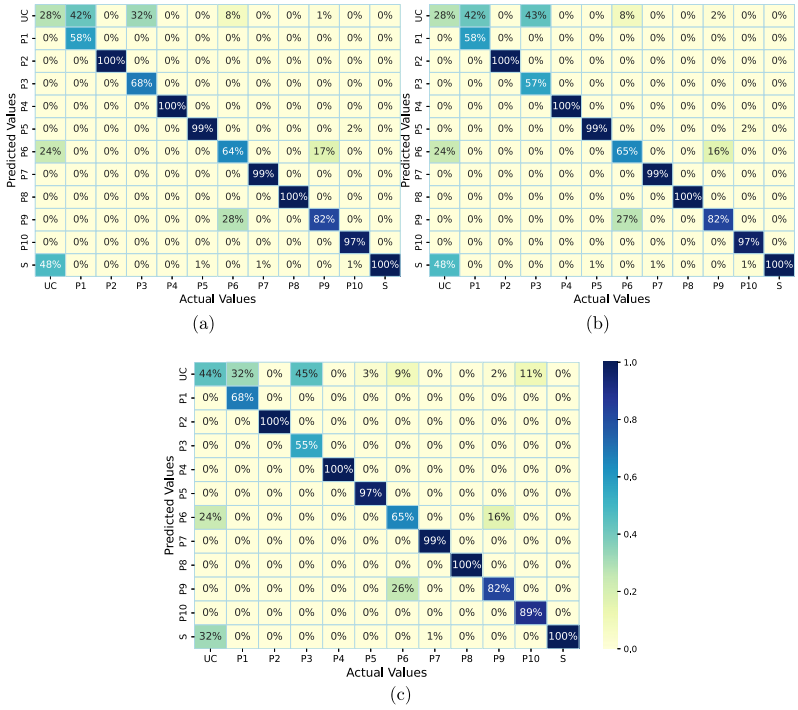



Figure 14. Confusion matrix; (a): SID, (b): SID-SAM, and (c): SID-SCM.

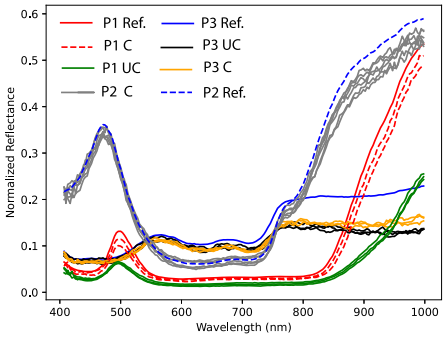



Figure 15. Spectra for pigment P1 and P3; solid red line (P1 Ref.) and solid blue line (P3 Ref.) are reference spectra for P1 and P3, respectively; red dashed line (P1 C) and solid green line (P1 UC) are spectra for classified and unclassified pixels of pigment P1; solid orange line (P3 C) and solid black line (P3 UC) are spectra for classified and unclassified pixels of pigment P3; dashed blue (P2 Ref.) and solid grey (P2 C) are spectra for reference and classified pixels of pigment 2.

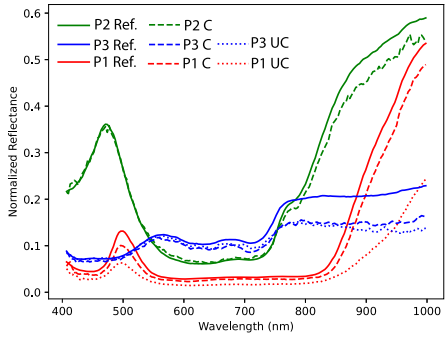



Figure 16. Spectrum of P1, P2, and P3; Ref., C and UC represent reference, classified and unclassified, respectively.

## APPENDIX A. REFLECTANCE SPECTRA OF 10 PIGMENTS AND SUBSTRATE USED TO TRAIN SVM, FC-NN AND 1D-CNN

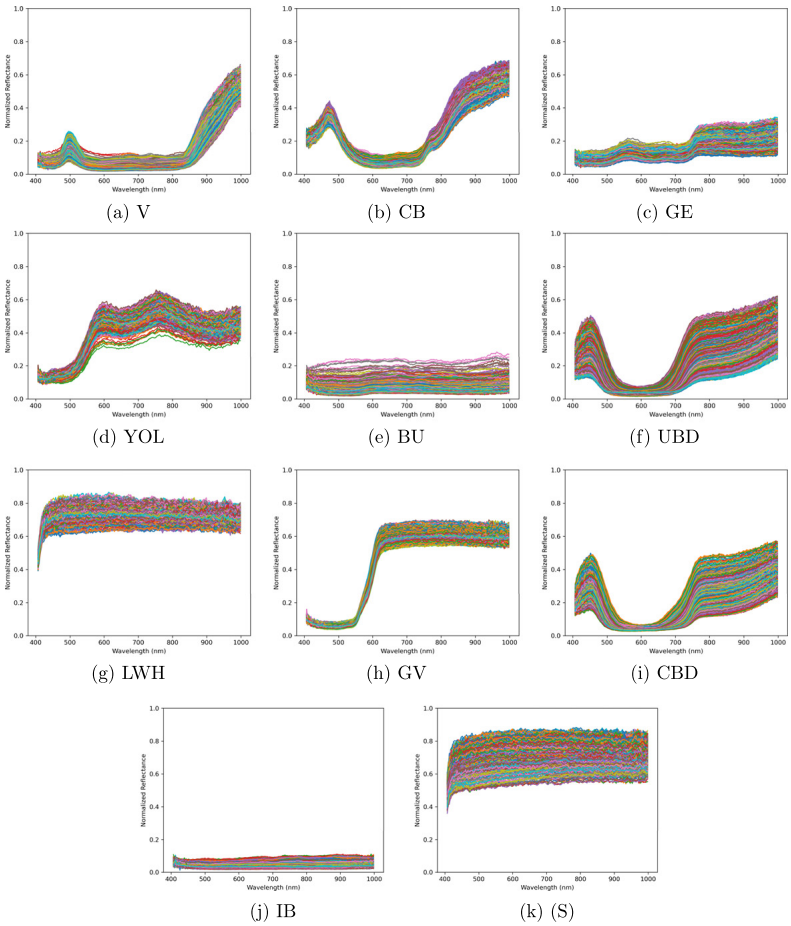
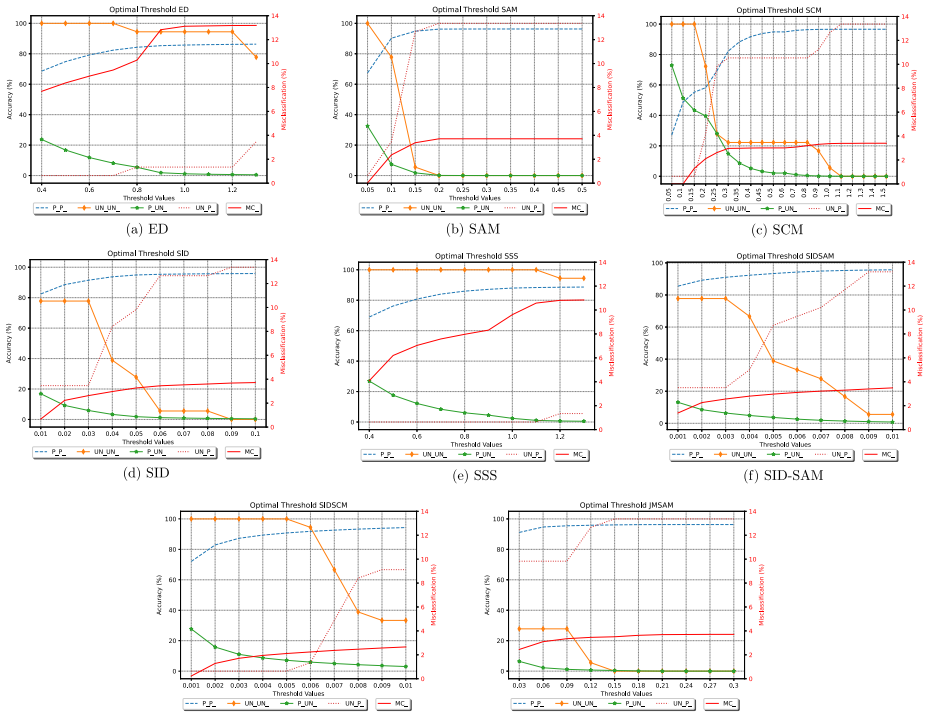



Figure A.1. Training spectra of 10 pigments and a substrate, plotted over spatial region of approximately 100 x 100 pixels with 186 spectral bands; Viridian (V), Cerulean Blue (CB), Green Earth (GE), Yellow Ochre Light (YOL), Burnt Umber (BU), Ultramarine Blue Deep (UBD), Lead White Hue (LWH), Genuine Vermilion (GV), Cobalt Blue Deep (CBD), Ivory Black (IB), and Substrate (S).

## APPENDIX B. GRAPH USED FOR DETERMINING THE OPTIMAL THRESHOLD VALUE FOR DIFFERENT ALGORITHMS




(a) ED (b) SAM (c) SCM

(d) SID (e) SSS (f) SID-SAM

(g) SID-SCM (h) (JMSAM)

Figure B.1. Classification accuracy graph of different algorithms at varying threshold values. The graph shows the accuracy of each algorithm in terms of pigment classified as a pigment (P_P_), unknown region classified as unknown (UN_UN_), pigment classified as unknown (P_UN_), unknown classified as a pigment (UN_P_) and pigment classifying as another pigment, i.e., misclassification (MC_).

## APPENDIX C. NORMALIZED REFLECTANCE SPECTRUM OF 10 PIGMENTS AND SUBSTRATE

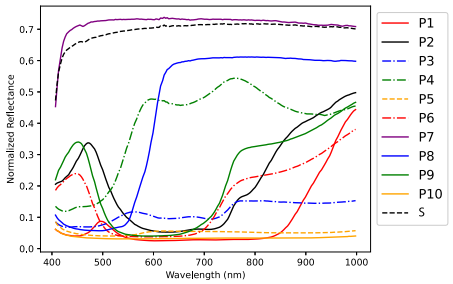



Figure C.1. Normalized reflectance spectrum for ten pigments and a substrate; P1: Veridian, P2: Cerulean Blue, P3: Green Earth, P4: Yellow Ochre Light, P5: Burnt Umber, P6: Ultramarine Blue Deep, P7: Lead White Hue, P8: Genuine Vermilion, P9: Cobalt Blue Deep, P10: Ivory Black, and S: Substrate.

## APPENDIX D. CLASSIFICATION RESULT FOR ALL USED ALGORITHMS

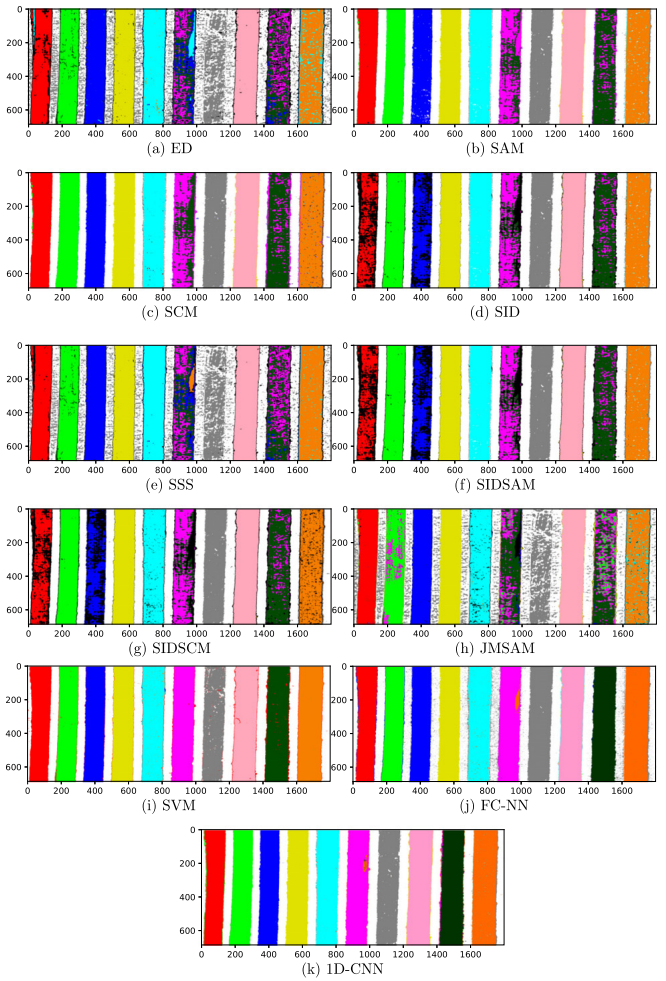



Figure D.1. Classification results obtained using various supervised and machine-learning algorithms; Euclidean Distance (ED), Spectral Angle Mapper (SAM), Spectral Correlation Mapper (SCM), Spectral Information Divergence (SID), Spectral Similarity Scale (SSS), Jeffries Matusita-Spectral Angle Mapper (JMSAM), Support Vector Machine (SVM), Fully Connected Neural Network (FC-NN) and One-dimensional Convolutional Neural Network (1D-CNN).

## REFERENCES


[1] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," IEEE Signal Process. Mag. **19**, 12–16 (2002).

[2] L. Tan and M.-l. Hou, "A study on the application of SAM classification algorithm in seal of calligraphy and painting based on hyperspectral technology," *2016 4th Int'l. Workshop on Earth Observation and Remote Sensing Applications (EORSA)* (IEEE, Piscataway, NJ, 2016), pp. 415–418.

[3] C. Balas, G. Epitropou, A. Tsapras, and N. Hadjinicolaou, "Hyperspectral imaging and spectral classification for pigment identification and mapping in paintings by el greco and his workshop," Multimedia Tools Appl. **77**, 9737–9751 (2018).

[4] M. K. Tripathi and H. Govil, "Evaluation of AVIRIS-NG hyperspectral images for mineral identification and mapping," Heliyon **5**, e02931 (2019).

[5] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," IEEE Trans. Geosci. Remote Sens. **42**, 1778–1790 (2004).

[6] N. Liu, Y. Guo, H. Jiang, and W. Yi, "Gastric cancer diagnosis using hyperspectral imaging with principal component analysis and spectral angle mapper," J. Biomed. Opt. **25**, 066005 (2020).

[7] L. Zhi, D. Zhang, J.-q. Yan, Q.-L. Li, and Q.-l. Tang, "Classification of hyperspectral medical tongue images for tongue diagnosis," Comput. Med. Imaging Graph. **31**, 672–678 (2007).

[8] B. Park, W. R. Windham, K. C. Lawrence, and D. P. Smith, "Contaminant classification of poultry hyperspectral imagery using a spectral angle mapper algorithm," Biosyst. Eng. **96**, 323–333 (2007).

[9] B. M. Devassy and S. George, "Contactless classification of strawberry using hyperspectral imaging," *CEUR Workshop Proc.* (CEUR-WS.org, Aachen, 2020).

[10] B. Park, W. R. Windham, K. C. Lawrence, and D. P. Smith, "Contaminant classification of poultry hyperspectral imagery using a spectral angle mapper algorithm," Biosyst. Eng. **96**, 323–333 (2007).

[11] D. Deepthi, B. M. Devassy, S. George, P. Nussbaum, and T. Thomas, "Classification of forensic hyperspectral paper data using hybrid spectral similarity algorithms," J. Chemom. **36**, e3387 (2022).

[12] J. van der Weerd, A. van Loon, and J. J. Boon, "Ftir studies of the effects of pigments on the aging of oil," Stud. Conserv. **50**, 3–22 (2005).

[13] A. Cosentino, "Fors spectral database of historical pigments in different binders," E-conserv. J. **2**, 57–68 (2014).

[14] T. Cavaleri, A. Giovagnoli, and M. Nervo, "Pigments and mixtures identification by visible reflectance spectroscopy," Procedia Chem. **8**, 45–54 (2013).

[15] D. Saunders and J. Kirby, "The effect of relative humidity on artists' pigments," Natl. Gallery Tech. Bull. **25**, 62–72 (2004) http://www.nationalgallery.org.uk/technical-bulletin/saunders_kirby2004.

[16] S. Lyu, X. Yang, N. Pan, M. Hou, W. Wu, M. Peng, and X. Zhao, "Spectral heat aging model to estimate the age of seals on painting and calligraphy," J. Cult. Herit. **46**, 119–130 (2020).

[17] J. L. Mass, R. Opila, B. Buckley, M. Cotte, J. Church, and A. Mehta, "The photodegradation of cadmium yellow paints in henri matisse's le bonheur de vivre (1905–1906)," Appl. Phys. A **111**, 59–68 (2013).

[18] C. Cuttle, "Damage to museum objects due to light exposure," Int. J. Ligh. Res. Technol. **28**, 1–9 (1996).

[19] L. Simonot and M. Elias, "Color change due to surface state modification," Color Res. Appl. **28**, 45–49 (2003) First published: 30 December 2002.

[20] P. P. Shinde and S. Shah, "A review of machine learning and deep learning applications," *2018 Fourth Int'l. Conf. on Computing Communication Control and Automation (ICCUBEA)* (IEEE, Piscataway, NJ, 2018), pp. 1–6.

[21] L. Cai, J. Gao, and D. Zhao, "A review of the application of deep learning in medical image classification and segmentation," Ann. Transl. Med. **8**, 713 (2020).

[22] L. Zhi, D. Zhang, J.-q. Yan, Q.-L. Li, and Q.-l. Tang, "Classification of hyperspectral medical tongue images for tongue diagnosis," Comput. Med. Imaging Graph. **31**, 672–678 (2007).

[23] B. R. Shivakumar and S. V. Rajashekararadhya, "Performance evaluation of spectral angle mapper and spectral correlation mapper classifiers over multiple remote sensor data," *2017 Second Int'l. Conf. on Electrical, Computer and Communication Technologies (ICECCT)* (IEEE, Piscataway, NJ, 2017), pp. 1–6.

[24] O. A. de Carvalho and P. R. Meneses, "Spectral correlation mapper (SCM): an improvement on the spectral angle mapper (SAM)," *Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18* (JPL Publication Pasadena, CA, 2000), Vol. 9.

[25] J. Qin, T. F. Burks, M. A. Ritenour, and W. G. Bonn, "Detection of citrus canker using hyperspectral reflectance imaging with spectral information divergence," J. Food Eng. **93**, 183–191 (2009).

[26] B. M. Devassy, S. George, and J. Y. Hardeberg, "Comparison of ink classification capabilities of classic hyperspectral similarity features," *2019 Int'l. Conf. on Document Analysis and Recognition Workshops (ICDARW)* (IEEE, Piscataway, NJ, 2019), Vol. 8, pp. 25–30.

[27] W. A. Malila, "Change vector analysis: An approach for detecting forest changes with landsat," *LARS Symposia* (IEEE, Piscataway, NJ, 1980), pp. 326–335.

[28] J. Chen, P. Gong, C. He, R. Pu, and P. Shi, "Land-use/land-cover change detection using improved change-vector analysis," Photogramm. Eng. Remote Sens. **69**, 369–379 (2003).

[29] O. de Carvalho Júnior, R. Guimarães, A. Gillespie, N. Silva, and R. Gomes, "A new approach to change vector analysis using distance and similarity measures," Remote Sens. **3**, 2473–2493 (2011).

[30] Y. Du, C.-I. Chang, H. Ren, C.-C. Chang, J. O. Jensen, and F. M. D'Amico, "New hyperspectral discrimination measure for spectral characterization," Opt. Eng. **43**, 1777–1786 (2004).

[31] M. N. Kumar, M. V. R. Seshasai, K. S. V. Prasad, V. Kamala, K. V. Ramana, R. S. Dwivedi, and P. S. Roy, "A new hybrid spectral similarity measure for discrimination among vigna species," Int. J. Remote Sens. **32**, 4041–4053 (2011).

[32] M. Zhang, Z. Qin, X. Liu, and S. L. Ustin, "Detection of stress in tomatoes induced by late blight disease in California, USA, using hyperspectral remote sensing," Int. J. Appl. Earth Obs. Geoinf. **4**, 295–310 (2003).

[33] H. Li, W. S. Lee, K. Wang, R. Ehsani, and C. Yang, "Extended spectral angle mapping (ESAM)' for citrus greening disease detection using airborne hyperspectral imaging," Precis. Agric. **15**, 162–183 (2014).

[34] M. Dabboor, S. Howell, M. Shokr, and J. Yackel, "The Jeffries-Matusita distance for the case of complex wishart distribution as a separability criterion for fully polarimetric SAR data," Int. J. Remote Sens. **35**, 6859–6873 (2014).

[35] S. Ullah, T. A. Groen, M. Schlerf, A. K. Skidmore, W. Nieuwenhuis, and C. Vaiphasa, "Using a genetic algorithm as an optimal band selector in the mid and thermal infrared (2.5–14 μm) to discriminate vegetation species," Sensors **12**, 8755–8769 (2012).

[36] S. Venkataraman, H. Bjerke, K. Copenhaver, and J. Glaser, "Optimal band selection of hyperspectral data for transgenic corn identification," *MAPPS/ASPRS 2006 Fall Conf.* (ASPRS, Baton Rouge, LA, 2006), pp. 6–10.

[37] L. Bruzzone, F. Roli, and S. B. Serpico, "An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection," IEEE Trans. Geosci. Remote Sens. **33**, 1318–1321 (1995).

[38] H. Deborah, N. Richard, and J. Y. Hardeberg, "On the quality evaluation of spectral image processing algorithms," *2014 Tenth Int'l. Conf. on Signal-Image Technology and Internet-Based Systems* (IEEE, Piscataway, NJ, 2014), pp. 133–140.

[39] J. Romero, A. Garcia-Beltrán, and J. Hernández-Andrés, "Linear bases for representation of natural and artificial illuminants," J. Opt. Soc. Am. A **14**, 1007–1014 (1997).

[40] H. Deborah, N. Richard, and J. Hardeberg, "A comprehensive evaluation of spectral distance functions and metrics for hyperspectral image processing," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. **8**, 3224–3234 (2015).

[41] KREMER Pigments, Books & Color Charts. Accessed: 6 October 2022.

[42] M. Herold, M. E. Gardner, and D. A. Roberts, "Spectral resolution requirements for mapping urban areas," IEEE Trans. Geosci. Remote Sens. **41**, 1907–1919 (2003).

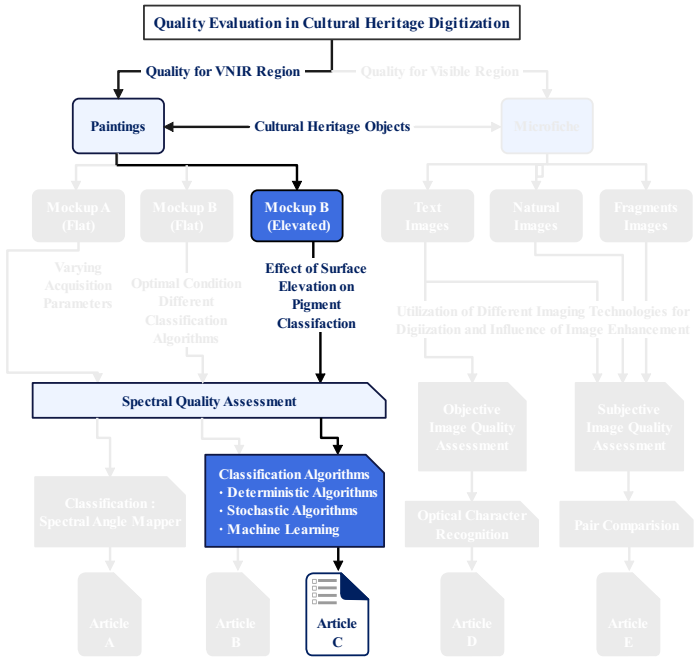
[43] X. Miao, P. Gong, S. Swope, R. Pu, and R. Carruthers, "Detection of yellow starthistle through band selection and feature extraction from hyperspectral imagery," Photogramm. Eng. Remote Sens. **73**, 1005–1015 (2007) .

[44] T. Mewes, J. Franke, and G. Menz, "Spectral requirements on airborne hyperspectral remote sensing data for wheat disease detection," Precis. Agric. **12**, 795–812 (2011).
[45] J. A. Gualtieri and S. Chettri, "Support vector machines for classification of hyperspectral data," *IGARSS 2000. IEEE 2000 Int'l. Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proc. (Cat. No. 00CH37120)* (IEEE, Piscataway, NJ, 2000), Vol. 2, pp. 813–815.
[46] E. Pouyet, T. Miteva, N. Rohani, and L. de Viguerie, "Artificial intelligence for pigment classification task in the short-wave infrared range," Sensors **21**, 6150 (2021).
[47] J. N. Sweet, "The spectral similarity scale and its application to the classification of hyperspectral remote sensing data," *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, 2003* (IEEE, Piscataway, NJ, 2003), pp. 92–99.
[48] F. A. Kruse, A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz, "The spectral image processing system (sips)—interactive visualization and analysis of imaging spectrometer data," Remote Sens. Environ. **44**, 145–163 (1993). Airbone Imaging Spectrometry.
[49] D. J. Mandal, S. George, M. Pedersen, and C. Boust, "Influence of acquisition parameters on pigment classification using hyperspectral imaging," J. Imaging Sci. Technol. **2021**, 334–346 (2021).
[50] R. N. Adep, A. P. Vijayan, A. Shetty, and H. Ramesh, "Performance evaluation of hyperspectral classification algorithms on AVIRIS mineral data," Perspectives Sci. **8**, 722–726 (2016). Recent Trends in Engineering and Material Sciences.
[51] C.-I. Chang, "Spectral information divergence for hyperspectral image analysis," *IEEE 1999 Int'l. Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No.99CH36293)* (IEEE, Piscataway, NJ, 1999), Vol. 1, pp. 509–511.
[52] H. Deborah, S. George, and J. Y. Hardeberg, "Pigment mapping of the scream (1893) based on hyperspectral imaging," *Int'l. Conf. on Image and Signal Processing* (Springer, Cham, 2014), pp. 247–256.
[53] E. Zhang, X. Zhang, S. Yang, and S. Wang, "Improving hyperspectral image classification using spectral information divergence," IEEE Geosci. Remote Sens. Lett. **11**, 249–253 (2013).
[54] J. C. Granahan and J. N. Sweet, "An evaluation of atmospheric correction techniques using the spectral similarity scale," *IGARSS 2001. Scanning the Present and Resolving the Future. Proc. IEEE 2001 Int'l. Geoscience and Remote Sensing Symposium (Cat. No. 01CH37217)* (IEEE, Piscataway, NJ, 2001), Vol. 5, pp. 2022–2024.
[55] L. Krauz, P. Páta, and J. Kaiser, "Assessing the spectral characteristics of dye-and pigment-based inkjet prints by VNIR hyperspectral imaging," Sensors **22**, 603 (2022).
[56] S. Padma and S. Sanjeevi, "Jeffries-Matusita-Spectral Angle Mapper (JM-SAM) spectral matching for species level mapping at Bhitarkanika, Muthupet and Pichavaram mangroves," Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **40**, 1403–1411 (2014).
[57] J. A. Gualtieri and S. Chettri, "Support vector machines for classification of hyperspectral data," *IGARSS 2000. IEEE 2000 Int'l. Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proc. (Cat. No. 00CH37120)* (IEEE, Piscataway, NJ, 2000), Vol. 2, pp. 813–815.
[58] M. A. Hamza, J. S. Alzahrani, A. Al-Rasheed, R. Alshahrani, M. Alamgeer, A. Motwakel, I. Yaseen, and M. I. Eldesouki, "Optimal and fully connected deep neural networks based classification model for unmanned aerial vehicle using hyperspectral remote sensing images," Can. J. Remote Sens. **48**, 681–693 (2022).
[59] F. Riese and S. Keller, "Soil texture classification with 1D convolutional neural networks based on hyperspectral data," ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. **IV-2/W5**, 615–621 (2019).
[60] C.-I. Chang, "An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis," IEEE Trans. Inf. Theory **46**, 1927–1932 (2000).
[61] J. Sweet, J. Granahan, and M. Sharp, "An objective standard for hyperspectral image quality," *Proc. AVIRIS Workshop* (JPL, Pasadena, CA, 2000).
[62] J. P. Kerekes, A. P. Cisz, and R. E. Simmons, "A comparative evaluation of spectral quality metrics for hyperspectral imagery," Proc. SPIE **5806**, 469–480 (2005).
[63] J. Zhang, Y. Zhang, and T. Zhou, "Classification of hyperspectral data using support vector machine," *Proc. 2001 Int'l. Conf. on Image Processing (Cat. No. 01CH37205)* (IEEE, Piscataway, NJ, 2001), Vol. 1, pp. 882–885.
[64] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," IEEE Trans. Geosci. Remote Sens. **42**, 1778–1790 (2004).
[65] S. Ding and L. Chen, "Classification of hyperspectral remote sensing images with support vector machines and particle swarm optimization," *2009 Int'l. Conf. on Information Engineering and Computer Science* (IEEE, Piscataway, NJ, 2009), pp. 1–5.
[66] H. Jabbar and R. Z. Khan, "Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study)," Comput. Sci. Commun. Instrum. Devices **70** (2015).
[67] D. Bau, J.-Y. Zhu, H. Strobelt, A. Lapedriza, B. Zhou, and A. Torralba, "Understanding the role of individual units in a deep neural network," Proc. Natl. Acad. Sci. **117**, 30071–30078 (2020).
[68] A. G. Schwing and R. Urtasun, "Fully connected deep structured networks". arXiv preprint arXiv:1503.02351v1 [cs.CV] 9 Mar 2015 (2015).
[69] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *Int'l. Conf. on Medical Image Computing and Computer-assisted Intervention* (Springer, Cham, 2015), pp. 234–241.
[70] S. Yu, S. Jia, and C. Xu, "Convolutional neural networks for hyperspectral image classification," Neurocomputing **219**, 88–98 (2017).
[71] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," IEEE Trans. Geosci. Remote Sens. **59**, 5966–5978 (2021).
[72] W. Hu, Y. Huang, W. Li, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," J. Sensors **2015**, 1–12 (2015).
[73] ZECCHI. Accessed: 10 June 2020.
[74] Norsk Elektro Optikk. Accessed: 20 December 2020.
[75] Spectralon multi-step targets. Accessed: 11 September 2020.
[76] Welcome to spectral python (SPy). Accessed: 08 August 2020.
[77] T. Fung and E. LeDrew, "Application of principal components analysis to change detection," Photogramm. Eng. Remote Sens. **53**, 1649–1658 (1987).
[78] T. Szandała, "Review comparison of commonly used activation functions for deep neural networks," *Bio-inspired Neurocomputing* (Springer, Cham, 2021), pp. 203–224.
[79] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization". arXiv preprint arXiv:1412.6980 (2014).
[80] M. Yaqub, J. Feng, M. S. Zia, K. Arshid, K. Jia, Z. U. Rehman, and A. Mehmood, "State-of-the-art CNN optimizer for brain tumor segmentation in magnetic resonance images," Brain Sci. **10**, 427 (2020).
[81] T. O'Malley, E. Bursztein, J. Long, F. Chollet, H. Jin, and L. Invernizzi, Kerastuner (2019) Accessed: 5 December 2022.

# Article C

## A Comparison of Pigment Classification Algorithms on Non–Flat Surfaces using Hyperspectral Imaging

D. J. Mandal, M. Pedersen, S. George, and C. Boust (2023). “Comparison of Pigment Classification Algorithms on Non–Flat Surfaces using Hyperspectral Imaging.” In: Manuscript under revision in Journal



# Comparison of Pigment Classification Algorithms on Non-Flat Surfaces using Hyperspectral Imaging

Dipendra J. Mandal[1], Marius Pedersen[1], Sony George[1], and Clotilde Boust[2]

[1]Department of Computer Science, Norwegian University of Science and Technology (NTNU), Norway
[2]Center for Research and Restoration of Museums of France (C2RMF), France

**Abstract**

Cultural heritage objects, such as paintings, provide valuable insights into the history and culture of human societies. Preserving these objects is of utmost importance, and developing new technologies for their analysis and conservation is crucial. Hyperspectral imaging is a technology with a wide range of applications in cultural heritage, including pigment classification. Pigment classification is crucial for conservators and curators in preserving works of art and acquiring valuable insights into the historical and cultural contexts associated with their origin. Various supervised algorithms, including machine learning, are used to classify pigments based on their spectral signatures. Since many artists employ impasto techniques in their artworks that produce a relief on the surface, i.e., transforming it from a flat object to a 2.5D or 3D, this further makes the classification task difficult. To our knowledge, no previous research has been conducted on pigment classification using hyperspectral imaging concerning an elevated surface. Therefore, this paper compares different spectral matching techniques that employ deterministic and stochastic methods, their hybrid combinations, and machine learning models for an elevated mockup to determine whether such topographical variation affects classification accuracy. In cultural heritage, the lack of adequate data is also a significant challenge for using machine learning, particularly in domains where data collection is expensive, time-consuming, or impractical. Data augmentation can help mitigate this challenge by generating new samples similar to the original. We also analyzed the impact of data augmentation techniques on the effectiveness of machine learning models for cultural heritage applications.

Keywords – *Pigment Classification, Impasto, Supervised Classification, Machine Learning, Hyperspectral Imaging*

# 1 Introduction

In Cultural Heritage (CH), paintings are an essential tangible component that provides valuable insights into our history, social norms, and beliefs. Therefore, the preservation and restoration of paintings are crucial and poses numerous challenges, including the removal of dirt and old varnish without damaging the paint layer [1] and the selection of appropriate materials for retouching [2]. To address these problems, it is essential to accurately identify the pigment used by the artist in an artwork. Scientific analysis sometimes required the physical samples, however, due to the nature of CH objects, it is not recommended to take samples from the object which in fact destroy the object even at a microscale and so very often, it is not permitted. Consequently, non-invasive or non-contact imaging techniques [3, 4, 5, 6] are necessary.

Hyperspectral Imaging (HSI) is a technology that has gained increasing attention in recent years due to its wide range of applications in various fields, including remote sensing [7], agriculture [8], medical sciences [9], forensics [10], biomedical engineering [11], and CH [12]. An important aspect of CH is pigment classification [13], and HSI can facilitate it by using spectral information about pigments and different classification algorithms. Artwork is not confined to two-dimensional canvases or boards, which means they are not always flat. The addition of relief, which introduces 2.5D or 3D to artwork, is also an important consideration [14, 15]. Several factors can contribute to this third dimension [16, 17, 18]; For example, morphological textures of brushstrokes on the painted surface [19], a thick layer of pigments applied by many renowned artists to their artwork for creating depth(impasto technique) [20, 21, 22, 23]. The geometry of a relief raised from a brush painting, impasto techniques, or any other factors may affect how light interacts with the surface, affecting the spectral signature captured by the hyperspectral sensor for a given pixel. Considering the importance of pigment identification or classification of an artwork, the influence of such factors must be explored.

Most of the research conducted so far has explored the effectiveness of various traditional supervised algorithms and machine learning models for pigment classification using HSI [24, 25]. However, those studies have primarily focused on flat surfaces and, to date, no research has investigated the same for artwork with an elevation. This paper aims to investigate how surface elevation in artworks affects the accuracy of pigment classification using HSI, with the underlying hypothesis that surface elevation impacts this accuracy. In CH, the lack of sufficient training datasets is also a considerable challenge for using machine learning, particularly in domains where collecting data is expensive, time-consuming, or impractical [26, 27]. Data augmentation is a technique that can help mitigate this challenge by generating new samples

similar to the original data. Therefore, in this paper, we have also compared and analyzed the impact of data augmentation techniques on the effectiveness of machine learning algorithms for pigment classification. This study is primarily concerned with answering the following research questions:

- To what extent does elevation of a surface in artworks affect the accuracy of pigment classification using hyperspectral imaging?
- What is the influence of data augmentation techniques on the efficacy of machine learning models for cultural heritage applications?

The rest of the paper is organized as follows. Section 2 briefly reviews the classification algorithms used in the CH field, focusing on pigment classification. Section 3 provides a brief overview of the classification algorithms used in this study. Section 4 describes the materials and methods used in the present study. Section 5 presents the results of the experiments and discusses the findings in detail. Finally, Section 6 concludes the paper and highlights directions for future research.

## 2 Background

Over the last few years, HSI, a non-invasive technique, has been widely employed for pigment classification in artwork [28, 5, 24]. It has resulted in significant advances in the study of spectral signatures and matching, broadening the scope of HSI technology in the CH domain. Molecules in the materials have unique vibration frequencies, which can be detected by analyzing how they absorb or reflect light at specific wavelengths. These characteristics of materials are known as spectral signatures and help to identify and distinguish pigment based on how it interacts with electromagnetic radiation. Many spectral matching algorithms have evolved in hyperspectral image processing, ranging from traditional clustering techniques to more recent automated matching models. The approaches used for matching spectra can be classified as deterministic, i.e., based on geometrical and physical aspects, or stochastic, which is based on the distributions [29]. These algorithms are essential for accurate and efficient pigment identification and analysis, making them a critical component of any HSI workflow. This section will briefly overview the classification algorithms employed in CH, specifically for pigment classification.

The Euclidean Distance (ED) metric is widely used to measure spectral similarity in HSI, and it works well when a data set has distinct or isolated clusters [30, 31, 32]. Mandal et al. [25] implemented ED and other supervised classification techniques to

classify pigments on a flat surface and observed that classification accuracy declines for some pigments with similar spectral characteristics.

Deborah et al. [5] explored the application of HSI in mapping the pigments of Edvard Munch's painting The Scream. They used two methods for spectral image classification, namely Spectral Angle Mapper (SAM) and Spectral Correlation Mapper (SCM). They observed that SCM performed better than SAM as it accounted for both positive and negative correlations between the spectra. Adjusting the threshold value for SAM reduced false detection, but it also decreased the accuracy of pigment classification, which varied depending on the type of pigment. They further suggested that different classes of pigments would require the use of distinct threshold values. In [33], the authors used the SAM technique to classify traditional Chinese pigments and recommended a similar suggestion of using different thresholds. The SAM algorithm measures the angle between two vectors, independent of the vector length, and therefore, insensitive to gain. As a result, this algorithm does not account for magnitude shifts in the spectrum. Please refer to Osmar et al. [34] for more information.

George and Hardeberg [35] demonstrate the usefulness of HSI to separate inks using SAM and SID as classification algorithms. They found that the SID algorithm perform better than SAM in cases where two distinct inks were overlaid. However, they also mentioned misclassification might arise from noise and non-uniformity in spectral signatures resulting from ink-paper blends. Mishra et al [36] utilized HSI to evaluate hybrid spectral similarity measures to classify paper samples used in forensic investigations. The findings indicate that the hybrid similarity measures of SIDSCM demonstrate better classification accuracy than conventional spectral similarity measures. Furthermore, the classification accuracies of SIDSCM and SIDSAM are comparable.

The authors in [37] discuss using SAM and Machine Learning (ML) models to classify and identify mineral pigments used in ancient Chinese paintings using HSI. The results show that, for similar colors and spectra, SAM is unable to classify; however, combining it with a decision tree can effectively improve the accuracy. The authors of [38] discuss the effectiveness of HSI technology in archaeological research for identifying and classifying materials in ancient tombs. They found that combining HSI data with Principal Component Analysis (PCA) transformation and SVM classification was an effective method for accurately classifying and identifying materials. The SVM classification based on feature bands improved classification accuracy and reduced data processing time. Kleynhans et al. [39] discuss using reflectance HSI and Neural Networks (NNs) to create labeled pigment maps of paintings. The authors report that a one-dimensional convolutional neural network (1D-CNN) model could accurately label the pigments in most of the paints studied. However, this find-

ing highlights the importance of having comprehensive training data for the model to perform well, the need for further studies to expand the training dataset, and the possibility of augmenting existing training datasets to develop a more robust solution.

Lie et al. [40] explore the potential of using NNs to analyze HSI data in the CH field. They present a thorough overview of the different applications and constraints of NNs models. The findings indicate that NNs offer a promising alternative to conventional statistical and multivariate analysis techniques for pigment identification and classification. The authors in [41] present a method for identifying pure pigments in CH using a combination of CNNs and SCM. The HSI data, collected within the range of 400nm to 720nm and at a resolution of 10nm, was pre-processed by smoothing and computing the first derivative before being fed to the network. The study emphasizes the significance of employing deep learning NNs for this application and the requirement for a comprehensive training dataset. A recent study by Mandal et al. [25] investigated the performance of various traditional supervised algorithms, their hybrid combinations, and ML models for pigment classification on flat surfaces. A research gap exists as the efficacy of these algorithms on non-flat objects has not been explored by any of the authors. Thus, further investigation is necessary for this area.

# 3 Classification Algorithms

This section presents the fundamental theory and mathematical expressions for the classification algorithms utilized in the study.

## 3.1 Euclidean Distance

Euclidean Distance (ED) is a distance metric that measures the distance between two points in an n-dimensional space [42]. It is calculated as the square root of the sum of the squared differences between the corresponding elements of the two points. In spectral analysis, the ED can compare the similarity between two spectra by measuring the difference between their respective pixel or spectral values. The formula for Euclidean distance between the image spectrum $t_i$ and a reference spectrum $r_i$, each with n elements, is defined using Equation (1).

$$ED = \sqrt{\sum_{i=1}^{nb} (t_i - r_i)^2} \tag{1}$$

*where, nb is the number of spectral bands.*

## 3.2 Spectral Angle Mapper

Spectral Angle Mapper (SAM), introduced by Boardman in 1992 [43], is a method for measuring the spectral similarity between two spectra, i.e., test and reference. This technique treats the spectra as $N$-dimensional vectors in space, where $N$ is the number of spectral bands and calculates the arccosine angles between them. The spectral angle, $\alpha$, between the two spectra is computed using Equation (2). A smaller angle indicates a better match between the spectra.

$$\alpha = \cos^{-1}\left(\frac{\sum_{i=1}^{nb} t_i r_i}{\sqrt{\sum_{i=1}^{nb} {t_i}^2}\sqrt{\sum_{i=1}^{nb} {r_i}^2}}\right) \tag{2}$$

*where $t_i$ represents the image spectrum, $r_i$ denotes the reference spectrum, and nb is the total number of bands.*

## 3.3 Spectral Correlation Mapper

Spectral Correlation Mapper (SCM) is one of several algorithms used in spectral similarity analysis for classification and feature extraction. It measures the Pearson correlation coefficient between two spectra by standardizing the data and centring them around the mean of the test and reference spectra. The result is then expressed as an angle using the arccosine function. This algorithm excludes negative correlation and retains the shading effect minimization characteristics similar to SAM, resulting in better classification results [34, 44]. SCM is computed using Equation (3).

$$\alpha = \cos^{-1}\left(\frac{\sum_{i=1}^{nb} (t_i - \bar{t}_i)(r_i - \bar{r}_i)}{\sqrt{\sum_{i=1}^{nb} (t_i - \bar{t}_i)^2 \sum_{i=1}^{nb} (r_i - \bar{r}_i)^2}}\right) \tag{3}$$

*where $\alpha$ is the arccosine of the spectral correlation measure in radians, $t_i$ and $\bar{t}_i$ are the image spectrum and its sample mean, similarly $r_i$ and $\bar{r}_i$ are the reference spectrum and its sample mean; and nb is the total number of bands.*

### 3.4 Spectral Information Divergence

In terms of spectral similarity, Spectral Information Divergence (SID) measures the dissimilarity between two spectra by comparing their spectral information content. It is based on the concept of Kullback-Leibler (KL) divergence, a measure of the difference between two probability distributions. In SID, the spectral information content of each pixel is modeled as a probability distribution, and the divergence between the two distributions is calculated. If two pixels have similar spectral information, their probability distributions will be similar, and the SID value will be low, and vice versa. The probability distribution of the test and reference spectra is expressed as Equation (4) and Equation (5), respectively [45].

$$p_i = \frac{t_i}{\sum_{i=1}^{nb} t_i} \tag{4}$$

$$q_i = \frac{r_i}{\sum_{i=1}^{nb} r_i} \tag{5}$$

where, $t_i$ is the image spectrum, $r_i$ is the reference spectrum, and $nb$ is the total number of bands. Using these two probability distributions, SID can be calculated with Equation (6).

$$SID = \sum_{i=1}^{nb} p_i \log\left(\frac{p_i}{q_i}\right) + \sum_{i=1}^{nb} q_i \log\left(\frac{q_i}{p_i}\right) \tag{6}$$

### 3.5 Spectral Similarity Scale

The Spectral Similarity Scale (SSS) provides a quantitative measure of the similarity between two spectra. This algorithm uses the Euclidean distance metric for magnitude and correlation to compare the shape of the spectra. This method combines both calculations, giving each equal weighting [46]. An SSS value of 0 indicates that the two spectra are identical, while a value of 1 indicates that the two spectra are entirely dissimilar. SSS is computed using Equation (7).

$$SSS = \sqrt{(d_e)^2 + (\hat{r})^2} \tag{7}$$

where, $d_e$ is the Euclidean distance between two spectra and is computed using

Equation (8) and its value ranges from 0 to 1 due to the factor $1/nb$.

$$d_e = \sqrt{\frac{1}{nb}\sum_{i=1}^{nb}(t_i - r_i)^2} \tag{8}$$

Equation (9) computes the value for $\hat{r}$, where r is the correlation coefficient between the two spectra and is computed using Equation (10).

$$\hat{r} = (1 - r^2) \tag{9}$$

$$r^2 = \left( \frac{\sum_{i=1}^{nb}(t_i - \bar{t}_i)(r_i - \bar{r}_i)}{\sqrt{\sum_{i=1}^{nb}(t_i - \bar{t}_i)^2 \sum_{i=1}^{nb}(r_i - \bar{r}_i)^2}} \right)^2 \tag{10}$$

## 3.6 Spectral information divergence-spectral angle mapper

The Spectral Information Divergence Spectral Angle Mapper (SIDSAM) is a hybrid approach that incorporates quantitative and qualitative matching measures. It utilizes the SID algorithm to assess the dissimilarity between two spectra and the SAM algorithm to evaluate their geometric similarity. This hybrid computation enhances the comparability of similar spectra and makes dissimilar spectra more distinctive, thus improving spectral discriminability. SIDSAM is computed by multiplying the SID by the tangent or sine function of the SAM, which calculates the perpendicular distance between the test and reference vectors. Both measures yield similar results, as reported in previous studies [47]. This hybrid computation SIDSAM can be calculated using either Equation (11) or (12).

$$SIDSAM = SID * \tan(SAM) \tag{11}$$

$$SIDSAM = SID * \sin(SAM) \tag{12}$$

where, SID and SAM are calculated using Equations (6) and (2), respectively.

### 3.7 Spectral information divergence-spectral correlation mapper

The Spectral Information Divergence Spectral Correlation Mapper (SIDSCM) is another hybrid approach that combines qualitative and quantitative metrics to increase spectral discriminability. It combines SID and SCM algorithms, similar to SIDSAM, where SID measures the difference between two spectra and SCM determines the Pearson correlation coefficient. To integrate the two measures, the product of SID and either the tangent or sine function of the correlation coefficient between two spectra is utilized [48]. The resultant method may be calculated using either Equations (13) or (14).

$$SIDSCM = SID * \tan(SCM) \tag{13}$$

$$SIDSCM = SID * \sin(SCM) \tag{14}$$

where SID and SCM can be computed using Equations (6) and (3) respectively.

### 3.8 Jeffries Matusita-Spectral Angle Mapper

The Jeffries Matusita (JM) distance is a statistical metric considering the covariance of two spectral vectors. The SAM method computes the angle between two spectral vectors to determine their spectral similarity. Jeffries Matusita-Spectral Angle Mapper (JMSAM) is calculated by first calculating the JM distance and then converting it to an angle with the inverse cosine method. The angle obtained is then compared to a threshold value to assess whether the two spectra belong to the same class. It can increase spectral classification accuracy by considering the spectral similarity and statistical distance [49]. It can be computed using either Equation (15) or (16).

$$JMSAM = JMD * \tan(SAM) \tag{15}$$

$$JMSAM = JMD * \sin(SAM) \tag{16}$$

where, JMD is JM distance and is computed using Equation (17)

$$JMD = 2\left(1 - e^{-B}\right) \tag{17}$$

*Here B is the Bhattacharyya distance and is computed using Equation (18).*

$$B = \frac{1}{8}(\mu_t - \mu_r)^T \left[\frac{\sigma_t + \sigma_r}{2}\right]^{-1} (\mu_t - \mu_r) + \frac{1}{2} ln \left[\frac{\left|\frac{\sigma_t + \sigma_r}{2}\right|}{\sqrt{|\sigma_t||\sigma_r|}}\right] \quad (18)$$

*Where, $\mu_t$ and $\mu_r$ are the mean of the test and reference spectra, respectively; $\sigma_t$ and $\sigma_r$ are the covariance of the test and reference spectra, respectively.*

## 3.9 Support Vector Machine

Support Vector Machine (SVM) is a machine-learning algorithm used for classification and regression analysis [50]. SVM classification aims to find a hyperplane that separates the data into two classes with maximum margin. It can handle non-linearly separable data using a kernel trick that maps it into a higher-dimensional space. This algorithm involves data preprocessing, such as normalization, to ensure that the features are on the same scale. It selects the most relevant features for the classification task. After that, it trains the model by finding the optimal hyperplane that maximizes the margin using a cost function. The cost function penalizes misclassified data points and encourages the SVM to find the hyperplane that separates the data with the largest margin. Finally, the model is tested on a validation set or test data set to evaluate its performance [51, 52].

## 3.10 1D-CNN

A Neural Network (NN) is a machine learning algorithm inspired by the structure and function of the human brain. The basic building block of a neural network is the artificial neuron, which takes inputs and applies a transformation to produce an output [53]. The architecture of a NN can vary widely depending on the task and the data being used. A 1D CNN is a type of NN commonly used for processing one-dimensional data such as time series, audio signals, and text data [54, 55, 56]. It consists of one or more convolutional layers, a pooling layer, and fully connected layers. The convolutional layer applies convolution operations to the input sequence using a set of learnable filters or kernels. This generates a set of feature maps representing the convolutional layer's output. The pooling layer is typically used to reduce the dimensionality of feature maps while maintaining the most important information. It applies an aggregation function such as max or average pooling to extract the most relevant features from each feature map. This helps reduce the number of parameters in the model and prevent overfitting. The fully connected layer takes the output of the pooling layer, flattens it into a one-dimensional vector, and passes it through a set of fully connected neurons. The output of the fully connected layer is often fed into a softmax function to generate class probabilities [57]. Overall, the

1D-CNN architecture is designed to extract and learn discriminative features from one-dimensional data sequences, making it suitable for various applications. Figure 1 illustrates the general architecture of a 1D CNN for use with HSI data.

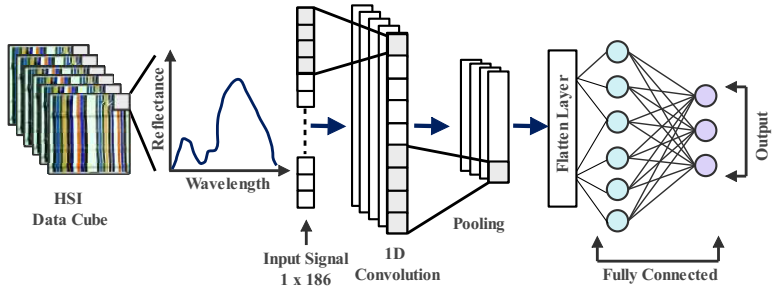



Figure 1: The architecture of 1D-CNN; typically comprises three fundamental layers: convolutional, pooling, and fully connected layer. Input from the HSI datacube (pixel value across wavelength) is fed to the convolutional layers, which apply a sliding window over given input data to perform feature extraction. The pooling layers reduce the size of the extracted features, and the fully connected layers classify the input based on the features obtained from the previous layers.

# 4 Materials and Methods

This section describes the mockup utilized in the study and the HSI acquisition laboratory setup. Additionally, we will describe the steps used to process the HSI data for the classification task.

## 4.1 Mockup

As shown in Figure 2, a pigment mockup was used and prepared in a laboratory environment. The mockup's base was constructed using 3D printing. It consisted of different elevation levels, including a flat surface and surfaces raised 2.5mm, 5mm, and 10mm from the base. A linen fabric was glued to the surface of the base. Three layers of white gesso were applied evenly across the entire surface of the canvas. The surface was carefully sanded between each layer using sandpaper to create a smooth and even coat.

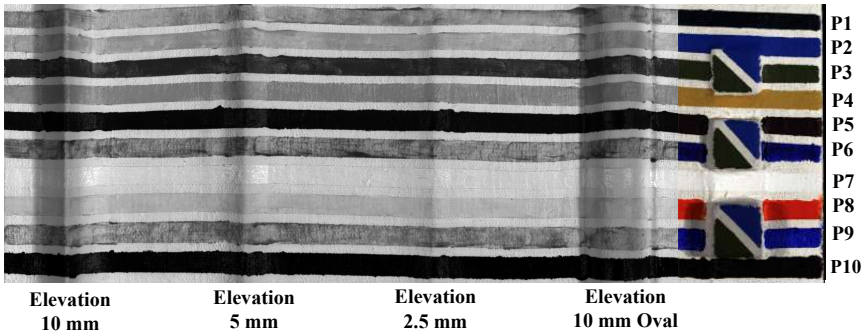



Figure 2: The pigment mockup used in the study comprised ten pigments labelled P1 to P10. These pigments were applied to a surface that included a flat region and three different elevation levels, namely 2.5mm, 5mm, and 10mm. For better visualization of the elevation, the left part of the image is shown in grayscale, captured at 998 nm in the IR region, while the right side is a colour image produced using bands at 640 nm, 551 nm, and 458 nm.

The selection of pigments for the research work was a critical step to ensure that the mockup accurately reflected the properties of pigments commonly used in historical artworks. The pigments were carefully chosen on the basis of their spectral characteristics and common appearance in CH research articles. The selection process also involved consultations with experts. Furthermore, web-based research was conducted on 164 known paintings from different centuries to determine the final selection of pigments. We selected the pigments that were used most frequently in those paintings. This approach ensured that the pigments used in the mockup represented those used in real paintings, making the research findings more applicable to real-world scenarios. The final selection of pigments included Veridian (V), Cerulean Blue (CB), Green Earth (GE), Yellow Ochre Light (YOL), Blue Cobalt (BC), Ultramarine Blue Deep (UBD), Lead White Hue (LWH), Genuine Vermilion (GV), Burnt Umber (BU), and Ivory Black (IB). Each pigment was applied to the mockup surface with a width of 6mm, with a 3mm gap between each pigment. The tubes were purchased from Zecchi, a reputable supplier of art materials [58]. Safflower oil was used for the whites, while linseed oil was used for all other pigment tubes.

## 4.2 HSI Acquisition Setup

Hyperspectral images were acquired in a laboratory using HySpex VNIR-1800, a line scanner camera developed by Norsk Electro Optikk [59], and a translation stage

setup, the schematic illustration of HIS is illustrated in Figure 3. The detector of the HySpex device consists of an actively cooled and stabilized complementary metal-oxide-semiconductor. It has a spectral range of 400 to 1000 nm with 186 spectral bands having a spectral resolution of 3.26 nm. A computer equipped with HySpex GROUND software provided by the manufacturing company controlled the hyperspectral acquisition system. This software automatically synchronizes the scanning speed for the defined integration time by the user. We used a close-range 30cm cylindrical lens for the acquisition. It captures 1800 spatial pixels across a line with a field of view of approximately 86mm.

During the experiment, we positioned the Spectralon® [60], a ColorChecker [61], and a pigment mockup on a movable part of the translation stage, as shown in Figure 3. These were placed at the same horizontal level and perpendicular to the camera's focal axis [62]. The Spectralon® used was a multi-step reference target with four adjacent panels with reflectance values of 99%, 50%, 25%, and 12%. It was used to calculate the normalized reflectance at the pixel level. To verify the obtained spectral data, we used a ColorChecker.

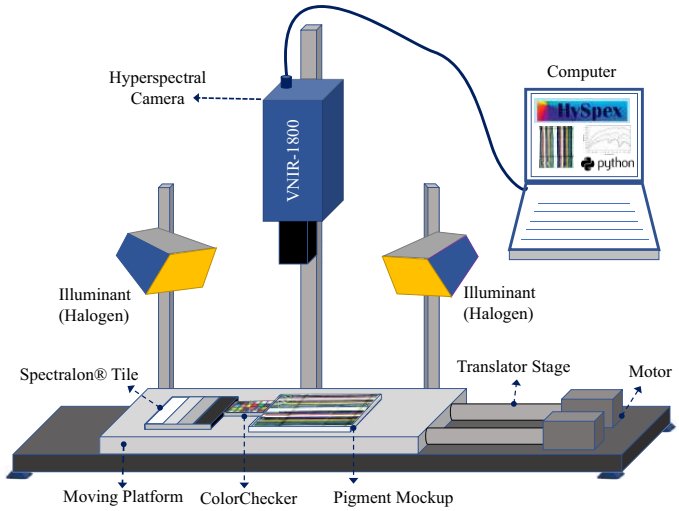



Figure 3: Schematic illustration of the HSI system utilized in the experiment. A 500 Watt halogen-based floodlight was used for illumination, and the illumination geometry was set to 45°-0°-45°, where 0° is the camera angle to normal.

## 4.3 Data Processing

The raw hyperspectral data are preprocessed for dark current factor, sensor correction, and radiometric calibration using HySpex RAD software provided by the camera manufacturer. The preprocessed data (converted to sensor-level absolute radiance value) are then converted to normalized reflectance using the known reflectance value of the reference target used in the experiment. The reference target surface might have some variation in the pixel value, so we averaged the values from 100 pixels for each line scan and calculated the reference target radiance value. Due to the small distance between the sensor and the object, we assumed that the path radiance effect was negligible. The spectral data were then cropped to exclude the ColorChecker and the reference target. The data processing steps were computed using the open-source software Python 3.9 [63]. Equation 19 provides the mathematical formulation used for conversion. Further data processing steps for supervised and ML-based classifications are discussed in the following sections.

$$R_{Obj}(\lambda) = R_{\mathrm{Ref\_t}}(\lambda)\frac{r_{Obj}(\lambda)}{r_{\mathrm{Ref\_t}}(\lambda)} \tag{19}$$

*where $R_{Obj}(\lambda)$ is the reflectance of an object, $R_{Ref\_t}(\lambda)$ is the reflectance of reference target, $r_{Obj}(\lambda)$ and $r_{Ref\_t}$ are the absolute sensor radiance values for the object and the reference target, respectively.*

### 4.3.1 Data Processing Steps for Supervised Classification Algorithms

To conduct supervised classification, reference or ground truth spectra are required to compare similarity. For this purpose, we selected a flat region with dimensions of around 10 × 10 pixels to establish a spectral library. We then saved the mean spectra for each pigment based on these regions, ensuring that the number of pixels was consistent. The plot for the spectrum of ten pigments and a substrate is included in the Appendix E. Our approach to constructing this library involved considering three different elevations for each pigment, as well as calculating an average spectrum that accounted for both elevated and flat surfaces. Data processing steps are shown with a block diagram in Figure 4.

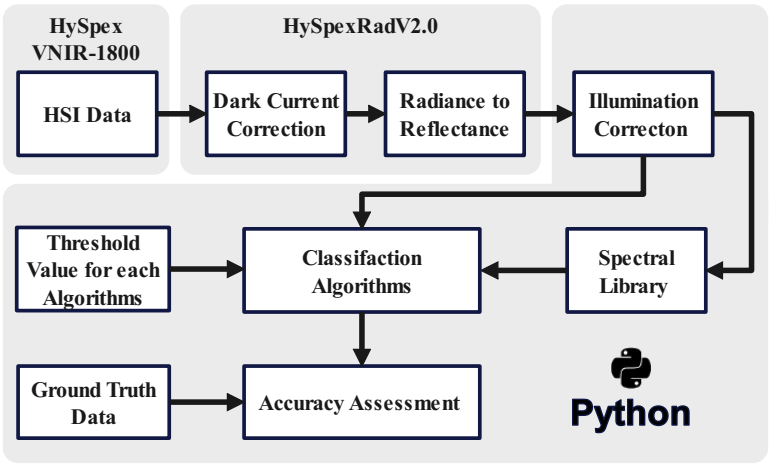



Figure 4: Workflow diagram illustrating the data processing Steps for supervised classification algorithms used in the study.

Selecting a threshold value is an important step in spectral matching [5, 64, 33], which involves identifying specific spectral ranges as belonging to one of several given classes. To achieve optimal classification results, the threshold value must balance minimizing misclassification rates and maximizing the number of correctly classified pigments. Mandal et al. [25] employed an empirical approach to determine an optimal threshold value, where the authors chose a small section of the HSI dataset from a mockup and extracted the reference spectrum by averaging 11 $\times$ 11 pixels from a flat region. They tested a range of values, computed the classification accuracy for various algorithms, and evaluated their accuracy using the confusion matrix. The threshold value used in our study was directly taken from this research and is mentioned in Table 1

### 4.3.2 Data Processing Steps for ML-based Classification Algorithms

Before feeding the normalized reflectance HSI data to the ML model, the data were labeled for distinct classes using a label encoder. We utilized one hot encoder for our dataset, where each class has one hot value (1), and the rest are cold (0). The data was then split into training and testing sets using an 80-20 split and was further standardized. Subsequently, the model was built and implemented. The training dataset was used to train the model by updating the weights and biases of neurons with each epoch until a considerably low Mean Square Error (MSE) and high accuracy

Table 1: The selected threshold value for eight different classification algorithms [25].

| Algorithms | Threshold Value |
|---|---|
| ED | 0.9 |
| SAM | 0.1 |
| SCM | 0.8 |
| SID | 0.03 |
| SSS | 1.1 |
| SIDSAM | 0.003 |
| SIDSCM | 0.005 |
| JMSAM | 0.09 |

were achieved. Once the model was trained, the test dataset was used to validate its performance. The overall workflow is illustrated in a block diagram in Figure 5.

Hyperparameter tuning is crucial in building robust and better generalized SVM models [65]. We tuned our SVM model for three key hyperparameters: kernel function, regularization (C), and gamma [51]. The kernel function transforms the input space into a higher-dimensional feature space, allowing the SVM to find a hyperplane (decision boundary) that can separate the classes. Several kernel functions are available, including linear, polynomial, and radial basis functions (RBF). The C parameter in SVM introduces a penalty for each misclassified data point. A smaller value of C results in a low penalty for misclassifications, leading to a decision boundary with a larger margin but more misclassifications. On the other hand, a larger value of C results in a higher penalty for misclassifications, leading to a decision boundary with a smaller margin and fewer misclassifications. The gamma parameter of RBF controls the distance of influence of a single training point. Low gamma values indicate a large similarity radius, resulting in more points being grouped together. For high gamma values, the points must be very close to each other to be considered in the same class. Therefore, models with very large gamma values tend to overfit. If the gamma is large, the effect of C becomes negligible. GridSearch cross-validation was used to optimize the hyperparameters of the SVM model [66]. This involved generating and testing the model for every possible combination of algorithm parameters specified in a grid. Table 2 shows the details of the hyperparameters.

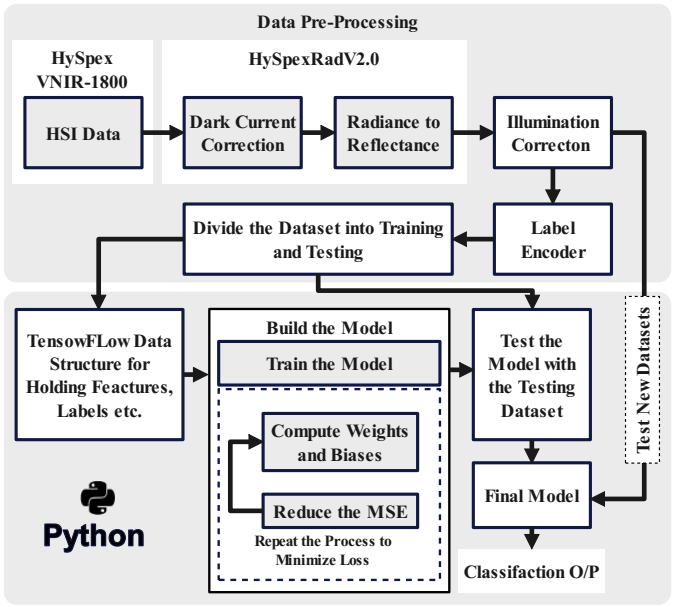



Figure 5: Workflow diagram illustrating the data processing steps for ML.

Table 2: SVM hyperparameters, tuning range, and the optimal value selected for classification.

| Hyperparameter | Range Used | Optimum Value Selected |
|---|---|---|
| Kernel | 'Polynomial', 'RBF', 'Sigmoid', 'Linear' | RBF |
| C | 0.1, 1, 10, 100, 1000 | 100 |
| Gamma | 1, 0.1, 0.01, 0.001,0.0001 | 1 |
| k-fold | 5 | 5 |

To optimize the hyperparameters of the 1D-CNN model, we utilized KerasTuner [**?**].The tuning process involved adjusting the number of convolutional layers, filter size, dropout rate, dense layer filter size, learning rate, and epoch. The resulting optimized model is illustrated in Figure 6, along with the specific hyperparameters used. The Adam optimizer with a learning rate of 0.001 and categorical cross-entropy loss function was used in the training process.

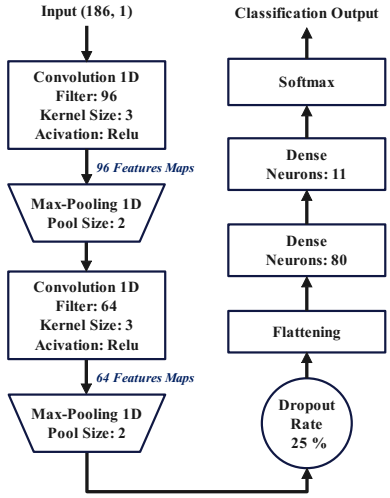



Figure 6: The architecture of a tuned 1D CNN model with optimized hyperparameters.

## 4.4 Data Augmentation

In ML, data augmentation is a method employed to expand the size of the training dataset by implementing various transformations on the available training data samples. The fundamental idea behind data augmentation is that alterations made to the labeled data should not modify the semantic interpretation of the labels [67, 68]. McFee et al. [69] suggested using deformation techniques that preserve the semantics of audio signals, improving the model's accuracy for the music classification task. Bjerrum et al. [70] used data augmentation techniques on spectral data to employ deep learning algorithms to predict drug composition in tablets using near-infrared regions. The results showed that data augmentation improves overall performance.

The data augmentation technique should be chosen based on the specific characteristics of the analyzed signal. Some techniques might be more appropriate than others, depending on the context. In our study, we augmented the datasets by introducing four attributes to the spectrum: offset, multiplication, Gaussian noise [71], and speckle noise [72]. An example of the implementation of these attributes is shown in Figure 7. Offset was varied $\pm$(0.0001 to 0.1 with a step size of 0.001) times the standard deviation of the training set. Multiplication was done with 1$\pm$(0.0001 to 0.1 with a step size of 0.001) times the standard deviation of the training set, and the two

different noises, Gaussian distributed additive noise and speckle which is a multiplicative noise, were added ten times with variation of 0.00001 and .000001 respectively. Using augmentation techniques, we produced two different training datasets. The first training dataset was generated by considering only a single spectrum from the spectral library, whereas for the second training dataset, we augmented each spectrum within the training dataset.

### 4.5 Accuracy Assessment

Accuracy assessment is an important step in evaluating the performance of classification algorithms. The most common and widely accepted method to express classification accuracy is confusion matrices. It helps to visualize the cross-tabulation of classified pigments; the matrix's main diagonal represents the correctly classified values, while the other elements indicate how many pixels in one category are incorrectly classified into other categories. For additional information on the confusion matrix, we refer to the work of Congalton [73]. For each algorithm, we calculate the accuracy for the predefined region illustrated in Figure 8.

## 5 Results and Discussion

This section presents the classification results obtained, along with an evaluation of the overall accuracy of the algorithms used, considering various reference spectra (ground truth). Furthermore, the outcomes of the SVM model employing data augmentation will also be elaborated.

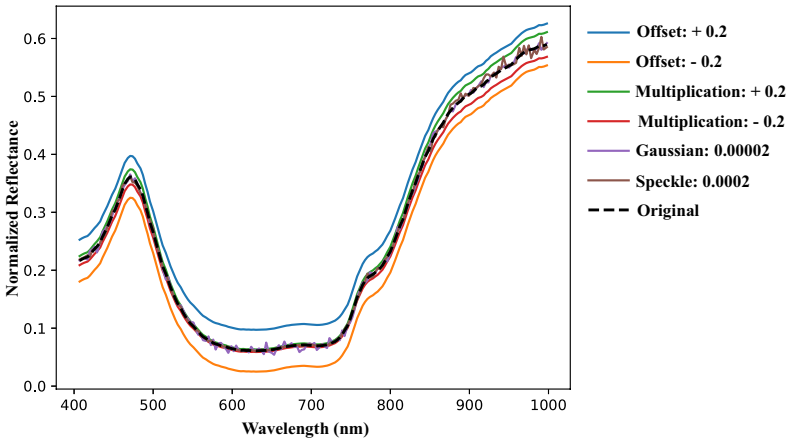



Figure 7: An example of data augmentation; is created with an addition of offset and multiplication to data by a factor of 0.2; Gaussian noise and speckle noise are added to the data with variations of 0.00002 and 0.0002, respectively.

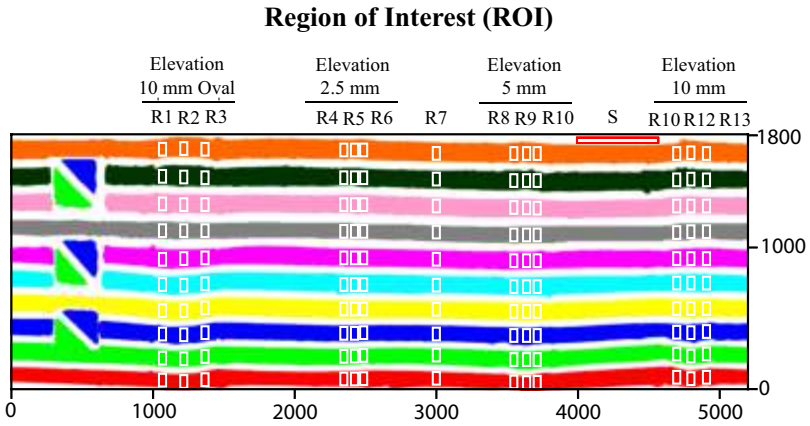



Figure 8: The ground truth image of the mockup showing regions of interest (ROIs) used for accuracy assessment corresponds to different elevation labels. R7 represents the flat region, while S represents the region for a substrate.

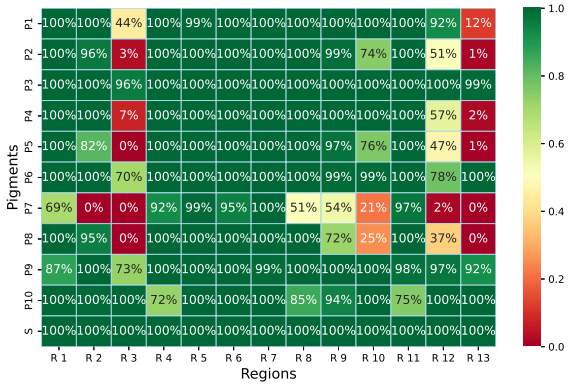
| | Flat Region | | | | | | | | | | Elevation : 2.5 mm | | | | | | | | | | Elevation : 5 mm | | | | | | | | | | Elevation : 10 mm | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| **ED** | 1 | 0.9 | 1 | 1 | 1 | 0.1 | 0.7 | 1 | 0.3 | 0.8 | 0.8 | 0.7 | 1 | 0.8 | 0.9 | 0.3 | 0.3 | 0.8 | 0.6 | 0.3 | 0.5 | 0.4 | 0.8 | 0.4 | 0.8 | 0.2 | 0.2 | 0.5 | 0.4 | 0.2 | 0.5 | 0.3 | 0.5 | 0.4 | 0.5 | 0.2 | 0.1 | 0.3 | 0.2 | 0.5 |
| | **77.29%** | | | | | | | | | | **66.24%** | | | | | | | | | | **44.48%** | | | | | | | | | | **35.55%** | | | | | | | | | |
| **SAM** | 1 | 1 | 1 | 1 | 1 | 0.2 | 1 | 1 | 0.6 | 1 | 0.9 | 1 | 0.9 | 1 | 0.9 | 0.7 | 0.9 | 1 | 0.9 | 0.5 | 0.9 | 1 | 0.8 | 1 | 0.7 | 0.6 | 0.6 | 1 | 0.9 | 0.3 | 1 | 1 | 0.9 | 1 | 0.9 | 0.9 | 0.5 | 1 | 0.7 | 0.5 |
| | **88.05%** | | | | | | | | | | **87.92%** | | | | | | | | | | **77.96%** | | | | | | | | | | **84.65%** | | | | | | | | | |
| **SCM** | 1 | 1 | 1 | 1 | 1 | 0.2 | 1 | 1 | 0.7 | 0.9 | 1 | 1 | 1 | 1 | 0.9 | 0.7 | 0.9 | 1 | 0.9 | 0.5 | 1 | 1 | 1 | 1 | 0.6 | 0.5 | 0.6 | 1 | 0.9 | 0.3 | 1 | 1 | 1 | 1 | 0.9 | 0.8 | 0.5 | 1 | 0.8 | 0.4 |
| | **88.08%** | | | | | | | | | | **89.89%** | | | | | | | | | | **79.33%** | | | | | | | | | | **83.91%** | | | | | | | | | |
| **SID** | 0.6 | 1 | 0.7 | 1 | 1 | 0.2 | 1 | 1 | 0.6 | 1 | 0.3 | 0.9 | 0.5 | 1 | 0.9 | 0.7 | 1 | 1 | 0.9 | 0.6 | 0.1 | 0.9 | 0.6 | 1 | 0.7 | 0.5 | 0.6 | 1 | 0.9 | 0.4 | 0.1 | 1 | 0.7 | 1 | 0.9 | 0.8 | 0.5 | 1 | 0.8 | 0.6 |
| | **80.86%** | | | | | | | | | | **76.77%** | | | | | | | | | | **66.73%** | | | | | | | | | | **72.87%** | | | | | | | | | |
| **SID-SAM** | 0.7 | 1 | 0.6 | 1 | 1 | 0.2 | 1 | 1 | 0.6 | 1 | 0.3 | 0.9 | 0.4 | 1 | 0.9 | 0.7 | 0.9 | 1 | 0.9 | 0.6 | 0.1 | 0.9 | 0.5 | 1 | 0.7 | 0.5 | 0.6 | 1 | 0.9 | 0.3 | 0.1 | 1 | 0.6 | 1 | 0.9 | 0.8 | 0.5 | 1 | 0.8 | 0.6 |
| | **79.98%** | | | | | | | | | | **76.09%** | | | | | | | | | | **66.01%** | | | | | | | | | | **72.22%** | | | | | | | | | |
| **SID-SCM** | 0.7 | 1 | 0.5 | 1 | 1 | 0.2 | 1 | 1 | 0.6 | 0.9 | 0.4 | 1 | 0.4 | 1 | 0.7 | 0.7 | 0.9 | 1 | 0.9 | 0.3 | 0.2 | 1 | 0.6 | 1 | 0.4 | 0.5 | 0.6 | 1 | 0.9 | 0.1 | 0.2 | 1 | 0.6 | 1 | 0.7 | 0.8 | 0.5 | 1 | 0.8 | 0.2 |
| | **79.86%** | | | | | | | | | | **74.16%** | | | | | | | | | | **62.59%** | | | | | | | | | | **67.68%** | | | | | | | | | |
| **JMSAM** | 1 | 0.6 | 1 | 1 | 1 | 0.1 | 0.7 | 1 | 0.3 | 0.8 | 0.9 | 0.6 | 1 | 1 | 0.6 | 0.4 | 0.4 | 1 | 0.7 | 0.3 | 0.7 | 0.4 | 1 | 1 | 0.4 | 0.3 | 0.5 | 1 | 0.7 | 0.2 | 0.8 | 0.5 | 0.9 | 1 | 0.3 | 0.6 | 0.3 | 0.9 | 0.7 | 0.3 |
| | **75.38%** | | | | | | | | | | **68.84%** | | | | | | | | | | **61.39%** | | | | | | | | | | **63.20%** | | | | | | | | | |
| **SSS** | 1 | 0.9 | 1 | 1 | 1 | 0.1 | 0.8 | 1 | 0.3 | 1 | 0.9 | 0.7 | 1 | 0.9 | 1 | 0.4 | 0.4 | 1 | 0.6 | 0.5 | 0.8 | 0.4 | 1 | 0.5 | 0.8 | 0.2 | 0.3 | 0.5 | 0.4 | 0.3 | 0.6 | 0.5 | 0.3 | 0.8 | 0.4 | 1 | 0.2 | 0.1 | 0.4 | 0.3 | 0.6 |
| | **80.03%** | | | | | | | | | | **74.05%** | | | | | | | | | | **52.30%** | | | | | | | | | | **44.86%** | | | | | | | | | |
| **SVM** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 0.9 | 1 | 1 | 0.9 | 1 | 0.4 | 0.7 | 1 | 0.9 | 0.7 | 0.5 | 1 | 0.5 | 0.5 | 0.9 | 0.3 | 0.5 | 1 | 0.9 |
| | **99.89%** | | | | | | | | | | **98.60%** | | | | | | | | | | **88.23%** | | | | | | | | | | **67.88%** | | | | | | | | | |
| **1D-CNN** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 1 | 1 | 0.6 | 1 | 1 | 0.8 | 1 | 1 | 1 | 1 | 1 | 1 | 0.4 | 1 | 1 | 0.9 |
| | **99.91%** | | | | | | | | | | **98.35%** | | | | | | | | | | **93.48%** | | | | | | | | | | **93.38%** | | | | | | | | | |

Figure 9: The classification accuracy from various algorithms used for evaluating ten different pigments on a flat surface and three elevated surfaces; the reference spectrum from the flat area was used for supervised and ML models; the color gradient utilized in this figure employs green to represent higher accuracy, red to indicate lower accuracy, and white represents an accuracy of 50%.

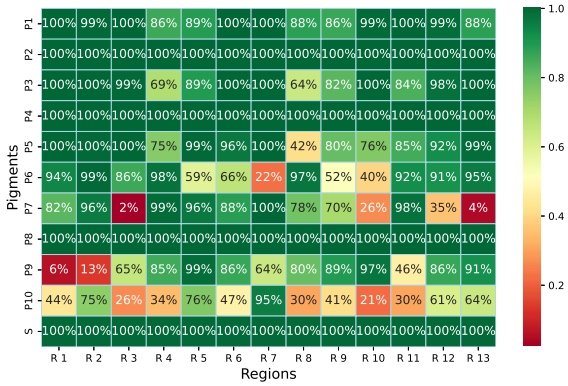
The Figure 9 illustrates the classification accuracy of ten different algorithms, including two machine learning models, for each pigment on both elevated surfaces and a flat surface. Algorithms 1D-CNN and SVM, outperformed all eight supervised algorithms, with the least accurate being algorithm ED. SAM and SCM performed better than the other algorithms after the machine learning models. On the other hand algorithms with hybrid approaches did not performed well overall.The images showing the classification results for each algorithm are attached in Appendix A.

Although there are variations in accuracy among the different algorithms, we can discern a pattern in the obtained classification results; the classification accuracy decreases with increasing elevation. An interesting observation is that SVM has lower accuracy for a region with an elevation of 10mm than most other algorithms for that same region. Figure 10a also presents the confusion matrix for SVM, revealing that the accuracy is particularly low for regions R3, R13, and R10, all located on one particular side of the elevation in the mockup (Figure 8). This side of the mockup has a shadow(Figure 2), which could be a significant factor in misclassifying these regions. Appendix B contains the confusion matrix for regions three and thirteen, illustrating the misclassification of pigments for SVM. The SAM algorithm determines the angle between two vectors, irrespective of their length, and thus its classification accuracy is less affected by any changes in the spectrum's magnitude. On the other hand, the SCM algorithm eliminates negative correlations while preserving the SAM characteristics. Therefore, these algorithms provide greater accuracy in the shadow area, as evidenced by the confusion matrix depicted in Figure 10b and Figure 10c for SAM and SCM, respectively.

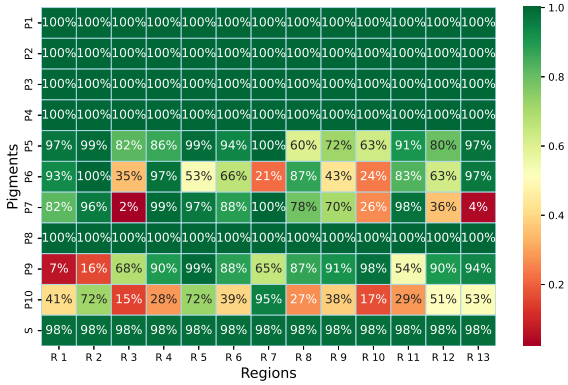
The accuracy of all supervised algorithms is lower for Pigment P6 and P9 in flat regions, but this is only due to the exact location chosen for accuracy assessment (R7), which includes most of the unclassified areas. This trend is not necessarily representative of all flat regions on average; also, elevated surfaces do not exhibit the same level of lower accuracy as flat regions. The classification results are presented in the Appendix A. However, the ED, JMSAM, and SSS algorithms have lower accuracy for Pigments P6 and P9 due to their similar spectra, which are also discussed by authors in [25]. This similarity is evident in Figure 11a, where the Pearson correlation coefficient between these two pigments is almost 1. similarly, Figure 11b also indicates a very low dissimilarity measure between these two pigments.




(a)



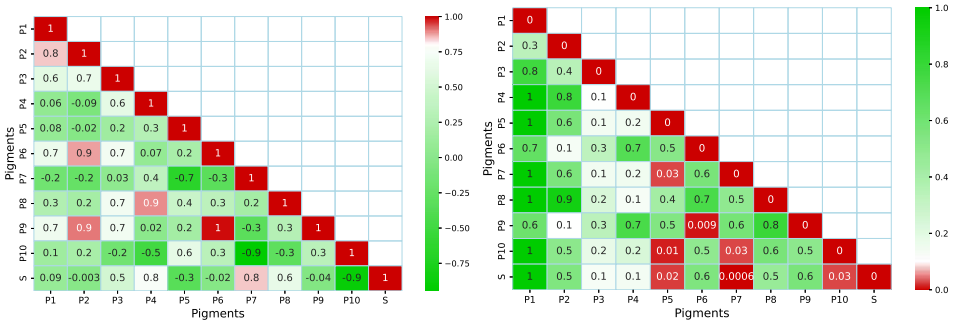

(b)




(c)



Figure 10: Classification accuracy for each pigment across all Regions of Interest (ROIs); For algorithms, (a): SVM, (b): SAM, and (c): SCM.

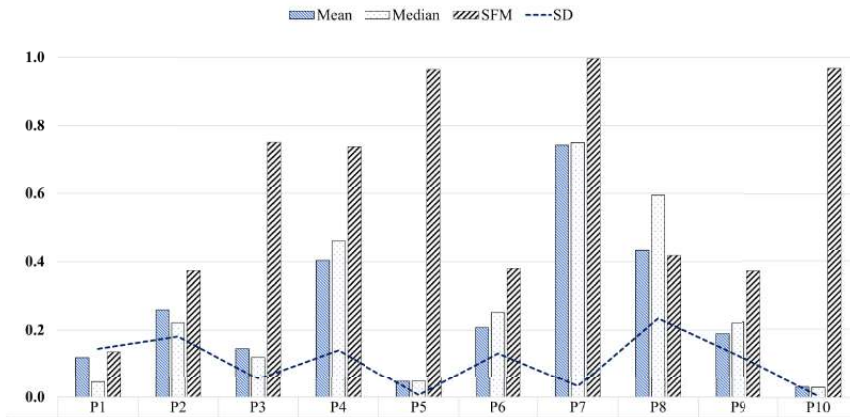



(a) (b)

Figure 11: The matrices display the correlation and dissimilarity between the pigments; (a) Pearson's correlation coefficient [74], a coefficient of 1 indicates a high correlation, while a value of zero represents no correlation; (b) SID calculated for dissimilarity, where a value of 0 implies a high degree of similarity between the spectra, while 1 indicates maximum dissimilarity.

The accuracy of the algorithms SID and its hybrid combination with SAM and SCM for pigment P1 has decreased as shown in Figure 9, and this decrease was further observed with an increase in elevation. It is important to note that not all pigments are affected in the same way. In the case of P1, the pixels were either classified as P1 or remained unclassified, as indicated by the confusion matrix for SID in the Appendix C. Mandal et al. [25] suggested that changing the threshold value could improve classification accuracy, but this approach could lead to misclassification of other pigments, which is generally undesirable in CH. The stochastic algorithm, SID, depends on the probability distributions of spectra, and alterations in data distribution can influence the overall entropy value. Moreover, the changes in the dataset can produce a varying impact on the entropy value for normal distribution and skewed distribution. In other words, if both distributions are shifted equally, the symmetric distribution would experience less change in entropy than the skewed distribution, owing to its higher predictability and lower uncertainty compared to the skewed distribution.

As illustrated in Figure 12, pigment P1 exhibits a notable difference between its mean and median values, resulting in a skewed distribution. Additionally, the standard deviation of this distribution is greater than its mean, indicating that the data points are more widely dispersed. Given the same amount of shift in datasets (identical in absolute terms), this shift may have a more significant impact on the dataset with a




| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Mean | 0.1192 | 0.2594 | 0.1454 | 0.4033 | 0.0462 | 0.2074 | 0.7422 | 0.4332 | 0.1894 | 0.0319 |
| Median | 0.0451 | 0.2220 | 0.1200 | 0.4601 | 0.0470 | 0.2511 | 0.7498 | 0.5944 | 0.2214 | 0.0306 |
| SFM | 0.136 | 0.374 | 0.751 | 0.738 | 0.964 | 0.38 | 0.996 | 0.418 | 0.37 | 0.968 |
| SD | 0.1448 | 0.1810 | 0.0552 | 0.1403 | 0.0063 | 0.1313 | 0.0322 | 0.2344 | 0.1243 | 0.0044 |

Figure 12: Statistical measures of mean, median, standard deviation (SD), and Spectral flatness measure (SFM), computed for all ten pigments.

larger spread than the one with data points being more tightly clustered around the mean, leading to a larger relative entropy between the two datasets. This is likely why some pigments are affected more than others. This also explains why we need to set different threshold values for different pigments.

The accuracy of Pigment P10 is higher for all the algorithms used on a flat surface, but it decreases significantly for supervised algorithms with elevation changes. On the other hand, ML models, SVM, and 1D-CNN show consistent classification accuracy for P10, regardless of elevation. When the reference spectrum is obtained from a flat region, it is likely to have minimal variation with test data sets within the same region, resulting in a higher number of correctly classified pixels. However, if the reference spectrum is taken from an elevated region, changes in reflectance values at higher elevations could cause higher variation and lead to more unclassified pixels. Based on this, one can hypothesize that using a reference spectrum taken from an elevated region is more likely to result in correct pixel classification on the same elevated surface region.

Low spectral intensity makes distinguishing between different land cover classes or features in an image difficult, leading to misclassification, especially for classes with

similar low reflectance values [75]. We observed that, in most cases, the P10 and P5 pigments were misclassified as each other or as the substrate, or they remained unclassified in most of the algorithms used. The Spectral Flatness Measure (SFM), also known as Wiener Entropy, is a metric that can be used to quantify the degree of flatness or peakiness of a spectrum by computing the geometric mean ratio to the arithmetic mean of the power spectrum [76]. We computed the SFM for the pigments, as shown in Figure 12. The SFM values for P7, P10, and P5 were higher, indicating a flat spectrum with nearly identical reflectance values across different wavelengths. This can make it difficult for a classifier algorithm to differentiate between different classes, resulting in lower classification accuracy.

Using a reference target with a surface height equivalent to the flat surface of the mockup, the normalized reflectance value was calculated for the HSI datacube. Ideally, the reflectance value of an elevated surface should be higher than that of a flat surface. However, shadows caused by the surface elevation lead to a decrease in the obtained reflectance value as the surface height increases. Generally, the substrate spectrum has a lower value than the P7 reference spectrum (the reference spectrum for ten pigments used and the substrate is provided in the Appendix E). With an increase in surface elevation, the distance between the P7 and its reference spectrum increases, while the distance between the P7 and substrate reference spectrum decreases. As a result, most regions for P7 are misclassified as substrate.

In Figure 13, we can see the overall classification accuracy obtained by averaging all pigments for four different surface elevations: flat, 2.5mm, 5mm, and 10mm, using all ten algorithms in the study. The reference spectrum used for building the spectral library and training the ML models was taken from the flat surface. For almost all algorithms, the classification accuracy for an elevated surface is lower compared to the flat region. Surfaces with a 2.5mm elevation have accuracy similar to or less than the flat surface, followed by 5mm and 10mm, respectively. This pattern is consistent for ED, SSS, SVM, and 1D-CNN algorithms. However, for the other six algorithms, we see that the accuracy at an elevation of 10mm is slightly greater than that of the surface with a 5mm elevation. The increase in overall classification accuracy can be attributed to the higher accuracy obtained for pigments P5 and P10. This higher accuracy for P5 and P10 might be due to the shadow effect; decreased reflectance value in the shadow region might have reduced the distance between the reference and measured pixels.

Earlier in this paper, we hypothesized that using a reference spectrum from an elevated surface would result in more accurate pixel classification for that same elevated surface region. To test this hypothesis, we built spectral libraries and training datasets using the reference spectrum from each elevated surface, i.e., 2.5mm, 5mm, and 10mm

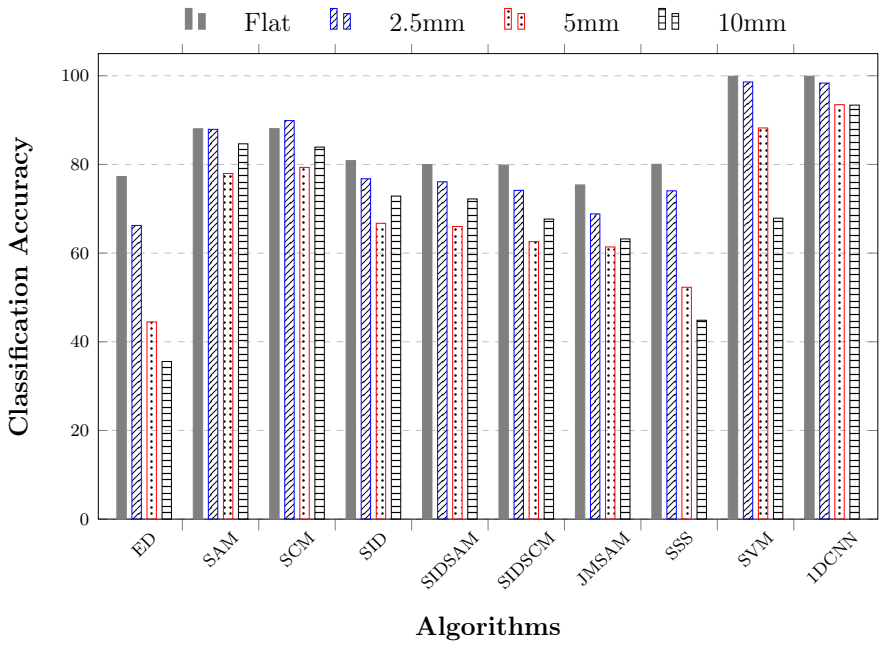



Figure 13: Overall classification accuracy for four different surface elevations (flat, 2.5mm, 5mm, and 10mm) using reference spectra from the flat surface.

and then computed the classification accuracy. In addition, we built spectral libraries and training datasets that represent an average spectrum of flat and varying elevation surfaces. To obtain these spectra, we averaged the pixel values over a spatial region taken as in a straight line (for example, in Figure 8, it is a pixels values of a line from right to left) for each band. The result from classifaction accuracy is summarized in Figure 14.

| | Ref. Spectrum: Flat | | | | Ref. Spectrum: 2.5mm | | | | Ref. Spectrum: 5mm | | | | Ref. Spectrum: 10mm | | | | Ref. Spectrum: Avg | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | F | E1 | E2 | E3 | F | E1 | E2 | E3 | F | E1 | E2 | E3 | F | E1 | E2 | E3 | F | E1 | E2 | E3 |
| **ED** | 0.77 | 0.66 | 0.44 | 0.36 | 0.79 | 0.75 | 0.51 | 0.42 | 0.71 | 0.7 | 0.48 | 0.42 | 0.53 | 0.58 | 0.53 | 0.44 | 0.82 | 0.7 | 0.48 | 0.41 |
| | **55.89%** | | | | **61.58%** | | | | **57.45%** | | | | **51.92%** | | | | **60.26%** | | | |
| **SAM** | 0.88 | 0.88 | 0.78 | 0.85 | 0.85 | 0.88 | 0.81 | 0.84 | 1 | 0.97 | 0.92 | 0.93 | 0.99 | 0.93 | 0.9 | 0.92 | 1 | 0.98 | 0.93 | 0.92 |
| | **84.65%** | | | | **84.52%** | | | | **95.31%** | | | | **93.50%** | | | | **95.75%** | | | |
| **SCM** | 0.88 | 0.9 | 0.79 | 0.84 | 0.86 | 0.87 | 0.78 | 0.82 | 0.91 | 0.95 | 0.91 | 0.92 | 0.98 | 0.93 | 0.89 | 0.91 | 0.99 | 0.96 | 0.88 | 0.91 |
| | **85.30%** | | | | **83.24%** | | | | **92.43%** | | | | **92.68%** | | | | **93.59%** | | | |
| **SID** | 0.81 | 0.77 | 0.67 | 0.73 | 0.75 | 0.81 | 0.73 | 0.72 | 0.85 | 0.88 | 0.8 | 0.82 | 0.84 | 0.83 | 0.8 | 0.86 | 0.95 | 0.92 | 0.83 | 0.88 |
| | **74.31%** | | | | **75.36%** | | | | **83.88%** | | | | **83.31%** | | | | **89.44%** | | | |
| **SID-SAM** | 0.8 | 0.76 | 0.66 | 0.72 | 0.75 | 0.8 | 0.72 | 0.71 | 0.83 | 0.88 | 0.79 | 0.82 | 0.79 | 0.81 | 0.78 | 0.84 | 0.94 | 0.92 | 0.83 | 0.87 |
| | **73.58%** | | | | **74.28%** | | | | **82.83%** | | | | **80.63%** | | | | **89.08%** | | | |
| **SID-SCM** | 0.8 | 0.74 | 0.63 | 0.68 | 0.74 | 0.77 | 0.66 | 0.68 | 0.74 | 0.84 | 0.8 | 0.8 | 0.74 | 0.79 | 0.78 | 0.83 | 0.89 | 0.89 | 0.79 | 0.86 |
| | **71.07%** | | | | **71.15%** | | | | **79.72%** | | | | **78.50%** | | | | **85.71%** | | | |
| **JMSAM** | 0.75 | 0.69 | 0.61 | 0.63 | 0.76 | 0.74 | 0.63 | 0.54 | 0.67 | 0.69 | 0.66 | 0.59 | 0.73 | 0.69 | 0.67 | 0.68 | 0.85 | 0.72 | 0.66 | 0.68 |
| | **67.20%** | | | | **66.58%** | | | | **65.19%** | | | | **69.14%** | | | | **72.76%** | | | |
| **SSS** | 0.8 | 0.74 | 0.52 | 0.45 | 0.83 | 0.84 | 0.59 | 0.57 | 0.73 | 0.76 | 0.59 | 0.49 | 0.7 | 0.67 | 0.62 | 0.58 | 0.91 | 0.78 | 0.57 | 0.56 |
| | **62.81%** | | | | **70.86%** | | | | **64.26%** | | | | **64.43%** | | | | **70.78%** | | | |
| **SVM** | 1 | 0.99 | 0.88 | 0.68 | 0.99 | 1 | 0.97 | 0.76 | 0.99 | 1 | 1 | 0.93 | 0.95 | 0.97 | 0.97 | 0.99 | 1 | 1 | 0.99 | 1 |
| | **88.65%** | | | | **93.12%** | | | | **97.89%** | | | | **97.24%** | | | | **99.70%** | | | |
| **1D-CNN** | 1 | 0.98 | 0.93 | 0.93 | 1 | 1 | 0.98 | 0.97 | 1 | 1 | 0.99 | 0.99 | 1 | 0.99 | 0.99 | 1 | 1 | 1 | 0.99 | 1 |
| | **96.28%** | | | | **98.62%** | | | | **99.42%** | | | | **99.35%** | | | | **99.49%** | | | |

Figure 14: Classifaction accuracy obtained using ten different algorithms for flat (F), 2.5 mm elevation (E1), 5 mm elevation(E2), and 10mm eleavation (E3), each computed for different spectral library and training datasets, built using reference spectrum from the flat region, three different elevated region and an average of all these.

Figure 14 shows that classification accuracy for the flat region (F) is higher when using reference spectra from the flat surface than at different elevations. Similarly, when using reference spectra from a 2.5mm elevated region, the classification accuracy is higher for regions with 2.5mm elevation (E1) for all algorithms. However, for regions with 5mm elevation (E2), the classification accuracy is almost identical to E1 and very close to other regions. In contrast, for regions with 10mm elevation (E3), the classification accuracy is greater than other surfaces, mainly for stochastic algorithms (SID, SIDSAM, SIDSCM), SVM, and 1D-CNN. However, it is the lowest for ED and SSS. The accuracy of SAM, SCM, and JMSAM is similar to E1 and E2 elevations but lower than the flat surface. Using average reference spectra improved accuracy for almost all flat and elevated regions compared to the accuracy obtained when using four different conditions of reference spectra. A detailed result displaying the classification

accuracy for individual pigments at each elevation and for different reference spectra conditions can be found in the Appendix F.

Result for average classification accuracy using ten algorithms and five reference spectra conditions (i.e., from the regions F, E1, E2, E3, and the average of these regions) is shown in Figure 15. When the reference spectrum was taken from the flat region, the classification accuracy was lower, or comparable (in the case of E1), for all algorithms compared to the accuracy obtained when the reference spectra were taken from regions E2, E3, or the average spectrum. The classification accuracy for most algorithms was almost the same when considering reference spectra from E2 or E3. Notably, using an average reference spectrum improved classification accuracy for almost all algorithms. The 1D-CNN algorithm had the highest classification accuracy among all the algorithms used, with slightly lower accuracy when the reference spectrum was taken from the flat region and almost similar results for all other reference spectra conditions.The classified images produced by the SVM and 1D-CNN algorithms using an average reference spectrum are provided in the Appendix D.

Data augmentation was performed to create additional training datasets for SVM, as shown in the flow diagram illustrated in Figure 16. Before computing classification accuracy, Hyperparameter tuning was performed using these augmented datasets. Finally, the SVM model was executed with the optimal hyperparameters of a polynomial kernel function and a regularization value of 0.1. Figure 17 shows the classification accuracy obtained for each pigment across all regions using the augmented training dataset where training data set from a flat region was selected. We observed that overall classification accuracy was higher than the SVM without data augmentation (Figure 10a). we also augmented the data from a single spectrum taken from the flat region first and then from the average spectrum; results for classification accuracy are included in Appendix G. Figure 18 shows the overall classification accuracy for each condition for flat and three different elevations.

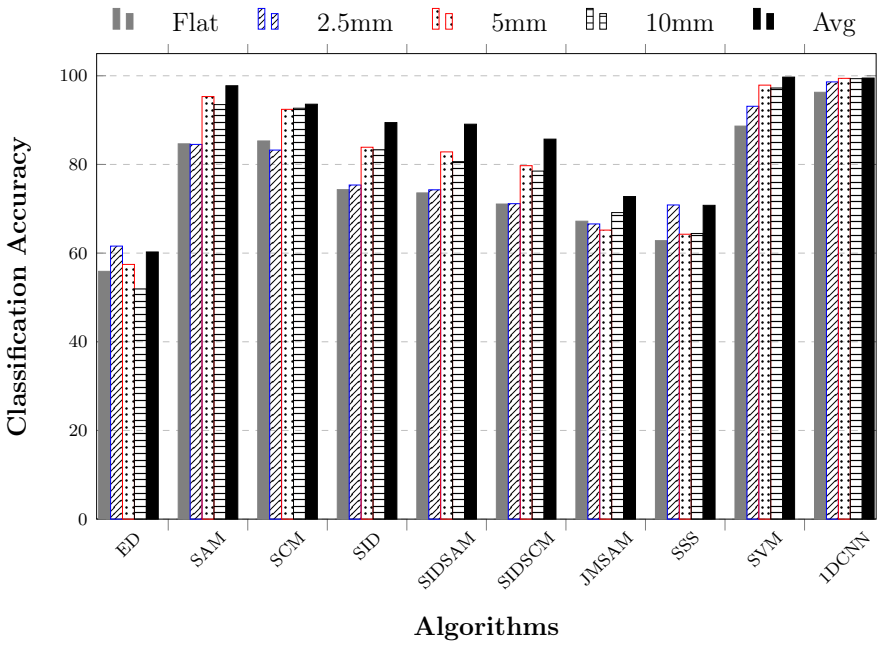



Figure 15: Overall classification accuracy for four surface elevations (flat, 2.5mm, 5mm, and 10mm) obtained by using reference spectra taken from an elevated region of 2.5 mm.

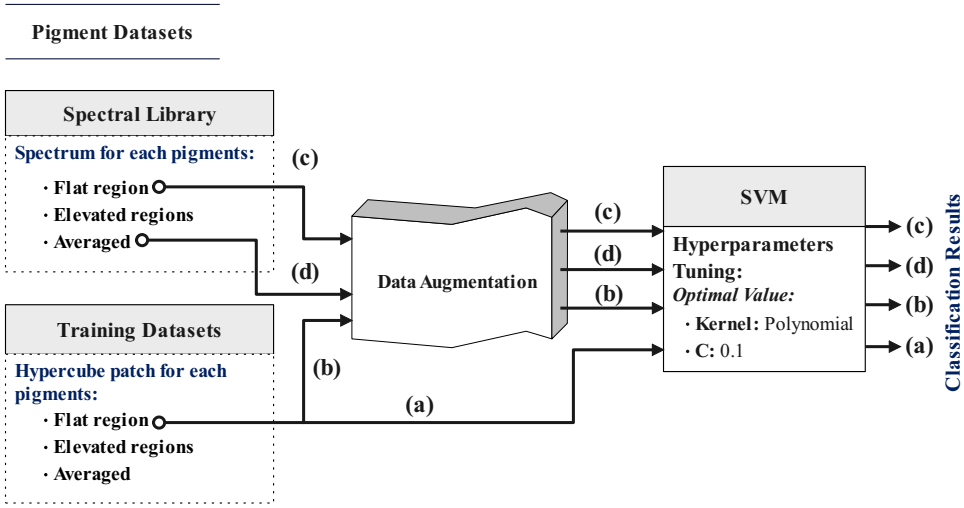



Figure 16: A workflow diagram illustrating the generation of training datasets using data augmentation for SVM; classification results from a, b, c and d are illustrated in Figure 19.

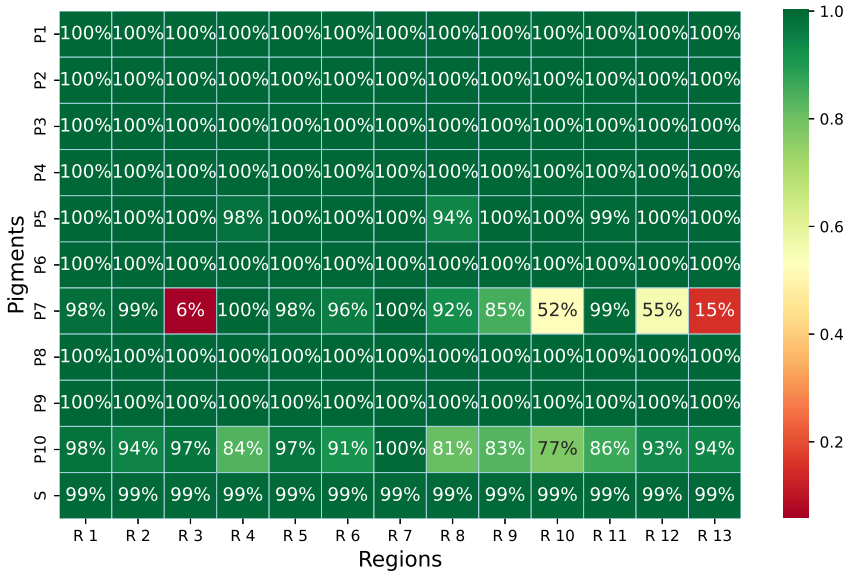



Figure 17: Classification accuracy for each pigment across all Regions of Interest (ROIs) for SVM using augmented dataset.

Classification accuracy for an augmented data set is higher when compared with all other three conditions, i.e., without data augmentation, data augmentation considering a single spectrum from a flat region and considering data augmentation considering a single average spectrum. However, using a single spectrum for augmentation yielded lower performance, except for an elevated region of 10 mm, where the accuracy was significantly improved compared to the non-augmented condition. Comparable accuracy was achieved using data augmentation with a flat region and augmenting a single averaged spectrum. As illustrated in Figure 9, SAM and SCM are the algorithms which performed better after ML models. Figure 18 shows that data augmentation using a single spectrum from a flat region is still better than SCM, highlighting that even with the single spectrum available, one can obtain better classification accuracy than supervised-based algorithms. The results suggest that data augmentation can improve classification accuracy, particularly when multiple spectra are augmented or when an average spectrum is used for augmentation. The classified images for each of these conditions are presented in Figure 19.

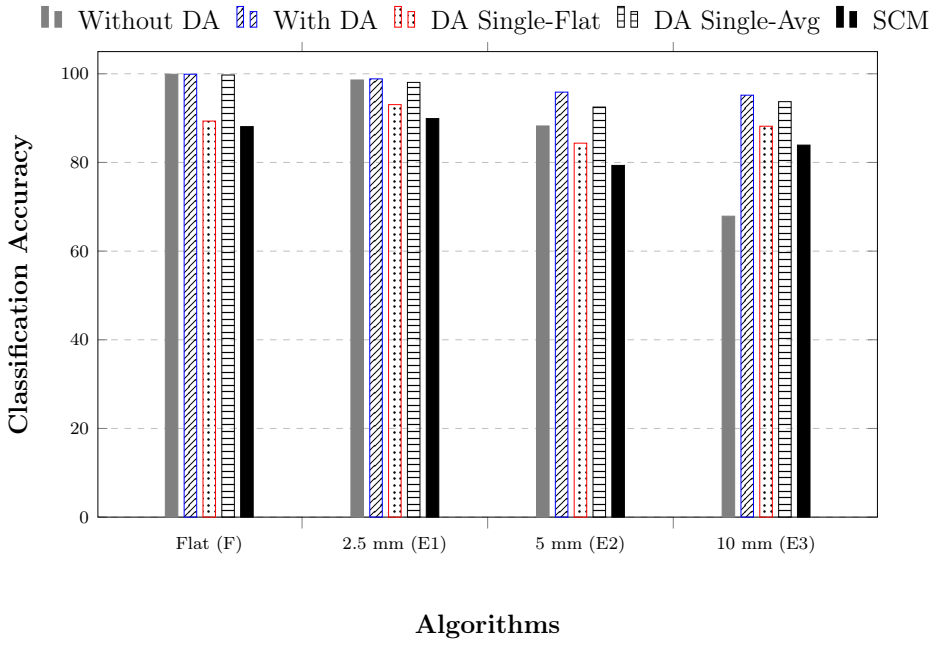



Figure 18: Overall classification accuracy for SVM at four different surface elevations (flat, 2.5 mm, 5 mm, and 10 mm) using different conditions of data augmentation, in comparison with SCM, employing reference spectra obtained from the flat region.

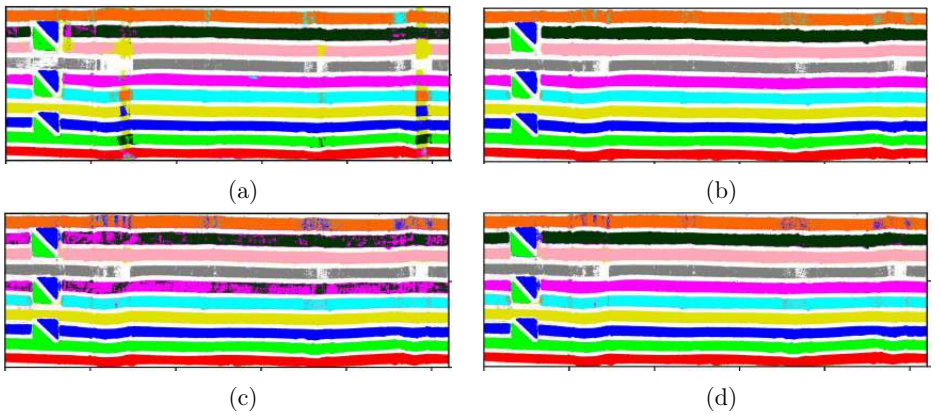



Figure 19: The classification results for SVM; (a): without data augmentation, (b): with data augmentation, (c): with data augmentation using single spectrum from flat region, and (d):with data augmentation using single spectrum from averaged region.

# 6 Conclusion

This study investigated the performance of different supervised algorithms and machine learning models for pigment classification using HSI on an elevated mockup. We have observed that the elevation itself does not significantly impact the classification accuracy; however, the elevation can result in the formation of shadows, which can have a significant effect on the classification accuracy of the algorithms used and varies for different algorithms. It was also observed that the choice of reference spectra plays a significant role in the accuracy of pigment classification. An average reference spectrum from different elevated regions yields better results than individual spectra. Among the ten algorithms tested, the 1D-CNN algorithm showed the highest classification accuracy, followed by SVM, SAM and SCM. Furthermore, results also indicated that data augmentation could significantly improve classification accuracy, particularly when multiple spectra are augmented or when an average spectrum is used for augmentation. This study provides valuable insights for analysing paintings in the CH domain. It could be beneficial in selecting appropriate classification algorithms when artworks have elevation or data that have shadows. In future work, removing shadows using image-processing techniques could be explored as a possible direction to further improve the accuracy of pigment classification.

# References

[1] Dusan Stulik, David Miller, Herant Khanjian, Narayan Khandekar, Janice Carlson, Richard Wolbers, and W Christian Petersen. *Solvent gels for the cleaning of works of art: the residue question*. Getty Publications, 2004.

[2] Shawn Digney-Peer, Karen Thomas, Roy Perry, Joyce Townsend, and Stephen Gritt. The imitative retouching of easel paintings. In *Conservation of easel paintings*, pages 626–653. Routledge, 2020.

[3] Raju Shrestha and Jon Yngve Hardeberg. Evaluation and comparison of multispectral imaging systems. In *Color and Imaging Conference*, volume 2014, pages 107–112. Society for Imaging Science and Technology, 2014.

[4] Costanza Cucci and Andrea Casini. Hyperspectral imaging for artworks investigation. In *Data handling in science and technology*, volume 32, pages 583–604. Elsevier, 2019.

[5] Hilda Deborah, Sony George, and Jon Yngve Hardeberg. Pigment mapping of the scream (1893) based on hyperspectral imaging. In *Image and Signal Processing:*

*6th International Conference, ICISP 2014, Cherbourg, France, June 30–July 2, 2014. Proceedings 6*, pages 247–256. Springer, 2014.

[6] Koen Janssens, Geert Van der Snickt, Frederik Vanmeert, Stijn Legrand, Gert Nuyts, Matthias Alfeld, Letizia Monico, Willemien Anaf, Wout De Nolf, Marc Vermeulen, et al. Non-invasive and non-destructive examination of artistic pigments, paints, and paintings by means of x-ray methods. *Analytical Chemistry for Cultural Heritage*, pages 77–128, 2017.

[7] Freek D. van der Meer, Harald M.A. van der Werff, Frank J.A. van Ruitenbeek, Chris A. Hecker, Wim H. Bakker, Marleen F. Noomen, Mark van der Meijde, E. John M. Carranza, J. Boudewijn de Smeth, and Tsehaie Woldai. Multi- and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation*, 14(1):112–128, 2012.

[8] Laura M Dale, André Thewis, Christelle Boudry, Ioan Rotar, Pierre Dardenne, Vincent Baeten, and Juan A Fernández Pierna. Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review. *Applied Spectroscopy Reviews*, 48(2):142–159, 2013.

[9] Guolan Lu and Baowei Fei. Medical hyperspectral imaging: a review. *Journal of Biomedical Optics*, 19(1):1 – 24, 2014.

[10] Gerda Edelman, Edurne Gaston, Ton van Leeuwen, P.J Cullen, and M.C.G Aalders. Hyperspectral imaging for non-contact analysis of forensic traces. *Forensic science international*, 223:28–39, 10 2012.

[11] Qingli Li, Xiaofu He, Yiting Wang, Hongying Liu, Dongrong Xu, and Fangmin Guo. Review of spectral imaging technology in biomedical engineering: achievements and challenges. *Journal of Biomedical Optics*, 18(10):1 – 29, 2013.

[12] Christian Fischer and Ioanna Kakoulli. Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications. *Studies in Conservation*, 51(sup1):3–16, 2006.

[13] W Stanley Taft and James W Mayer. The science of paintings. *Measurement Science and Technology*, 12(5):653–653, 2001.

[14] Peter Paul Rubens. *The Virgin as the Woman of the Apocalypse*. J. Paul Getty Museum. about 1623–1624.

[15] Joyce H Townsend. Painting techniques and materials of turner and other british artists 1775-1875. In *Historical painting techniques, materials, and studio practice: preprints of a symposium, University of Leiden, the Netherlands, 26-29 June 1995*, pages 176–185, 1995.

[16] Sylvana Barrett and Dusan C Stulik. An integrated approach for the study of painting techniques. In *Historical painting techniques, materials, and studio practice: preprints of a symposium, University of Leiden, the Netherlands, 26-29 June 1995*, pages 6–11, 1995.

[17] Eddie Sinclair. The polychromy of exeter and salisbury cathedrals: a preliminary comparison. In *Historical painting techniques, materials, and studio practice: preprints of a symposium, University of Leiden, the Netherlands, 26-29 June 1995*, pages 105–110. Getty Conservation Institute Malibu, 1995.

[18] William Baxter, Jeremy Wendt, and Ming C Lin. Impasto: A realistic, interactive model for paint. In *Proceedings of the 3rd International Symposium on Non-photorealistic Animation and Rendering*, pages 45–148, 2004.

[19] Yunfei Fu, Hongchuan Yu, Chih-Kuo Yeh, Jianjun Zhang, and Tong-Yee Lee. High relief from brush painting. *IEEE transactions on visualization and computer graphics*, 25(9):2763–2776, 2018.

[20] J Salvant Plisson, Laurence de Viguerie, L Tahroucht, M Menu, and G Ducouret. Rheology of white paints: How van gogh achieved his famous impasto. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 458:134–141, 2014.

[21] Karin Groen. Investigation of the use of the binding medium by rembrandt. *Zeitschrift fur Kunsttechnologie und Konservierung*, 2(2):208–211, 1997.

[22] W.S Elkhuizen, T.W.J Dore-Callewaert, Emilien Leonhardt, Abbie Vandivere, Y Song, S.C Pont, Jo M.P Geraedts, and J Dik. Comparison of three 3d scanning techniques for paintings, as applied to vermeer's 'girl with a pearl earring. *Heritage science*, 7(1):1–22, 2019.

[23] Victor Gonzalez, Marine Cotte, Gilles Wallez, Annelies van Loon, Wout De Nolf, Myriam Eveno, Katrien Keune, Petria Noble, and Joris Dik. Unraveling the composition of rembrandt's impasto through the identification of unusual plumbonacrite by multimodal x-ray diffraction analysis. *Angewandte Chemie*, 131(17):5675–5678, 2019.

[24] Dipendra J Mandal, Sony George, Marius Pedersen, and Clotilde Boust. Influence of acquisition parameters on pigment classification using hyperspectral

imaging. In *Color and Imaging Conference*, volume 2021, pages 334–346. Society for Imaging Science and Technology, 2021.

[25] Dipendra J. Mandal, Marius Pedersen, Sony George, Hilda Deborah, and Clotilde Boust. An experiment-based comparative analysis of pigment classification algorithms using hyperspectral imaging. Manuscript submitted for publication.

[26] Marco Fiorucci, Marina Khoroshiltseva, Massimiliano Pontil, Arianna Traviglia, Alessio Del Bue, and Stuart James. Machine learning for cultural heritage: A survey. *Pattern Recognition Letters*, 133:102–108, 2020.

[27] Stanislav Smirnov and Alma Eguizabal. Deep learning for object detection in fine-art paintings. In *2018 Metrology for Archaeology and Cultural Heritage (MetroArchaeo)*, pages 45–49. IEEE, 2018.

[28] Costas Balas, George Epitropou, Athanasios Tsapras, and Nicos Hadjinicolaou. Hyperspectral imaging and spectral classification for pigment identification and mapping in paintings by el greco and his workshop. *Multimedia Tools and Applications*, 77(8):9737–9751, 2018.

[29] Sennaraj Vishnu, Rama Rao Nidamanuri, and R Bremananth. Spectral material mapping using hyperspectral imagery: a review of spectral matching and library search methods. *Geocarto international*, 28(2):171–190, 2013.

[30] Nirmal Keshava. Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. *IEEE Transactions on Geoscience and remote sensing*, 42(7):1552–1565, 2004.

[31] Jianfeng Li, D Brynn Hibbert, Steven Fuller, Julie Cattle, and Christopher Pang Way. Comparison of spectra using a bayesian approach. an argument using oil spills as an example. *Analytical chemistry*, 77(2):639–644, 2005.

[32] Daniele Cerra, Jakub Bieniarz, Janja Avbelj, Rupert Müller, and Peter Reinartz. Spectral matching through data compression. volume Volume, pages 1–4, 01 2011.

[33] GONG Mengting and FENG Pingli. Preliminary study on the application of hyperspectral imaging in the classification of and identification chinese traditional pigments classification——a case study of spectral angle mapper. *Sciences of conservation and archaeology*, (4):76–83, 2014.

[34] O Abilio De Carvalho and Paulo Roberto Meneses. Spectral correlation mapper (SCM): an improvement on the spectral angle mapper (SAM). In *Summaries of*

*the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18*, volume 9. JPL Publication Pasadena, CA, 2000.

[35] Sony George and Jon Yngve Hardeberg. Ink classification and visualisation of historical manuscripts: Application of hyperspectral imaging. In *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, pages 1131–1135, 2015.

[36] Deepthi, Binu Melit Devassy, Sony George, Peter Nussbaum, and Tessamma Thomas. Classification of forensic hyperspectral paper data using hybrid spectral similarity algorithms. *Journal of Chemometrics*, 36(1):e3387, 2022.

[37] Cheng Fan, Pengchang Zhang, Shuang Wang, and Bingliang Hu. A study on classification of mineral pigments based on spectral angle mapper and decision tree. In *Tenth International Conference on Digital Image Processing (ICDIP 2018)*, volume 10806, pages 1639–1643. SPIE, 2018.

[38] Mingyan Gu, Shuqiang Lyu, Miaole Hou, Sheng Ma, Zhenhua Gao, Shuzhang Bai, and Pingping Zhou. Classification and recognition of tomb information in hyperspectral image. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives*, 42(3):411–416, 2018.

[39] Tania Kleynhans, Catherine M Schmidt Patterson, Kathryn A Dooley, David W Messinger, and John K Delaney. An alternative approach to mapping pigments in paintings with hyperspectral reflectance image cubes using artificial intelligence. *Heritage Science*, 8(1):1–16, 2020.

[40] Lingxi Liu, Tsveta Miteva, Giovanni Delnevo, Silvia Mirri, Philippe Walter, Laurence de Viguerie, and Emeline Pouyet. Neural networks for hyperspectral imaging of historical paintings: A practical review. *Sensors*, 23(5):2419, 2023.

[41] Ailin Chen, Rui Jesus, and Marcia Vilarigues. Convolutional neural network-based pure paint pigment identification using hyperspectral images. In *ACM Multimedia Asia*, pages 1–7. 2021.

[42] John Clifford Gower. Properties of euclidean and non-euclidean distance matrices. *Linear algebra and its applications*, 67:81–97, 1985.

[43] F.A. Kruse, A.B. Lefkoff, J.W. Boardman, K.B. Heidebrecht, A.T. Shapiro, P.J. Barloon, and A.F.H. Goetz. The spectral image processing system (sips)—interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44(2):145–163, 1993. Airbone Imaging Spectrometry.

[44] Osmar de Carvalho Júnior, Renato Guimarães, Alan Gillespie, Nilton Silva, and Roberto Gomes. A new approach to change vector analysis using distance and similarity measures. *Remote Sensing, vol. 3, issue 11, pp. 2473-2493*, 3:2473–2493, 11 2011.

[45] Chein-I Chang. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis. *IEEE Transactions on Information Theory*, 46(5):1927–1932, 2000.

[46] John P Kerekes, Adam P Cisz, and Rulon E Simmons. A comparative evaluation of spectral quality metrics for hyperspectral imagery. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, volume 5806, pages 469–480. SPIE, 2005.

[47] Yingzi Du, Chein-I Chang, Hsuan Ren, Chein-Chi Chang, James O Jensen, and Francis M D'Amico. New hyperspectral discrimination measure for spectral characterization. *Optical engineering*, 43(8):1777–1786, 2004.

[48] M Naresh Kumar, MVR Seshasai, KS Vara Prasad, V Kamala, KV Ramana, RS Dwivedi, and PS Roy. A new hybrid spectral similarity measure for discrimination among vigna species. *International journal of remote sensing*, 32(14):4041–4053, 2011.

[49] S. Padma and S. Sanjeevi. Jeffries matusita-spectral angle mapper (jm-sam) spectral matching for species level mapping at bhitarkanika, muthupet and pichavaram mangroves. volume 40, page 1403 – 1411, 2014.

[50] Lipo Wang. *Support vector machines: theory and applications*, volume 177. Springer Science & Business Media, 2005.

[51] Bernhard E Boser, Isabelle M Guyon, and Vladimir N Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152, 1992.

[52] William S Noble. What is a support vector machine? *Nature biotechnology*, 24(12):1565–1567, 2006.

[53] Oludare Isaac Abiodun, Aman Jantan, Abiodun Esther Omolara, Kemi Victoria Dada, Nachaat AbdElatif Mohamed, and Humaira Arshad. State-of-the-art in artificial neural network applications: A survey. *Heliyon*, 4(11):e00938, 2018.

[54] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on*

*Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

[55] Shiqi Yu, Sen Jia, and Chunyan Xu. Convolutional neural networks for hyperspectral image classification. *Neurocomputing*, 219:88–98, 2017.

[56] Danfeng Hong, Lianru Gao, Jing Yao, Bing Zhang, Antonio Plaza, and Jocelyn Chanussot. Graph convolutional networks for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 59(7):5966–5978, 2021.

[57] Wei Hu, Yangyu Huang, Wei Li, Fan Zhang, and Hengchao Li. Deep convolutional neural networks for hyperspectral image classification. *Journal of Sensors*, 2015:1–12, 07 2015.

[58] ZECCHI, soluzioni per artisti, materiale per restauro. http://www.zecchi.it/. Accessed: 14 April 2021.

[59] Norsk Elektro Optikk. http://www.hyspex.no/. Accessed: 20 December 2020.

[60] Spectralon multi-step targets. https://www.labspherestore.com/product-p/aa-006xx-000.htm. Accessed: 11 September 2020.

[61] X-Rite ColorChecker Classic. https://www.xrite.com/. Accessed: 11 September 2020.

[62] David Saunders and John Cupitt. Image processing at the national gallery: The vasari project. 14:72–85, 1993.

[63] Python Software Foundation. Python Language Reference, Version 3.10. https://docs.python.org/3/reference/, 2021. [Accessed 12 December 2020].

[64] Li Tan et al. A study on the application of sam classification algorithm in seal of calligraphy and painting based on hyperspectral technology. In *2016 4th International Workshop on Earth Observation and Remote Sensing Applications (EORSA)*, pages 415–418. IEEE, 2016.

[65] Hilde JP Weerts, Andreas C Mueller, and Joaquin Vanschoren. Importance of tuning hyperparameters of machine learning algorithms. *arXiv preprint arXiv:2007.07588*, 2020.

[66] Andreas C Müller and Sarah Guido. *Introduction to machine learning with Python: a guide for data scientists*. " O'Reilly Media, Inc.", 2016.

[67] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019.

[68] Qingsong Wen, Liang Sun, Fan Yang, Xiaomin Song, Jingkun Gao, Xue Wang, and Huan Xu. Time series data augmentation for deep learning: A survey. *arXiv preprint arXiv:2002.12478*, 2020.

[69] Brian McFee, Eric J Humphrey, and Juan Pablo Bello. A software framework for musical data augmentation. In *ISMIR*, volume 2015, pages 248–254. Citeseer, 2015.

[70] Esben Jannik Bjerrum, Mads Glahder, and Thomas Skov. Data augmentation of spectral data for convolutional neural network (cnn) based deep chemometrics. *arXiv preprint arXiv:1710.01927*, 2017.

[71] Wilbur B Davenport, William L Root, et al. *An introduction to the theory of random signals and noise*, volume 159. McGraw-Hill New York, 1958.

[72] Ajay Kumar Boyat and Brijendra Kumar Joshi. A review paper: noise models in digital image processing. *arXiv preprint arXiv:1505.03489*, 2015.

[73] Russell G. Congalton. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1):35–46, 1991.

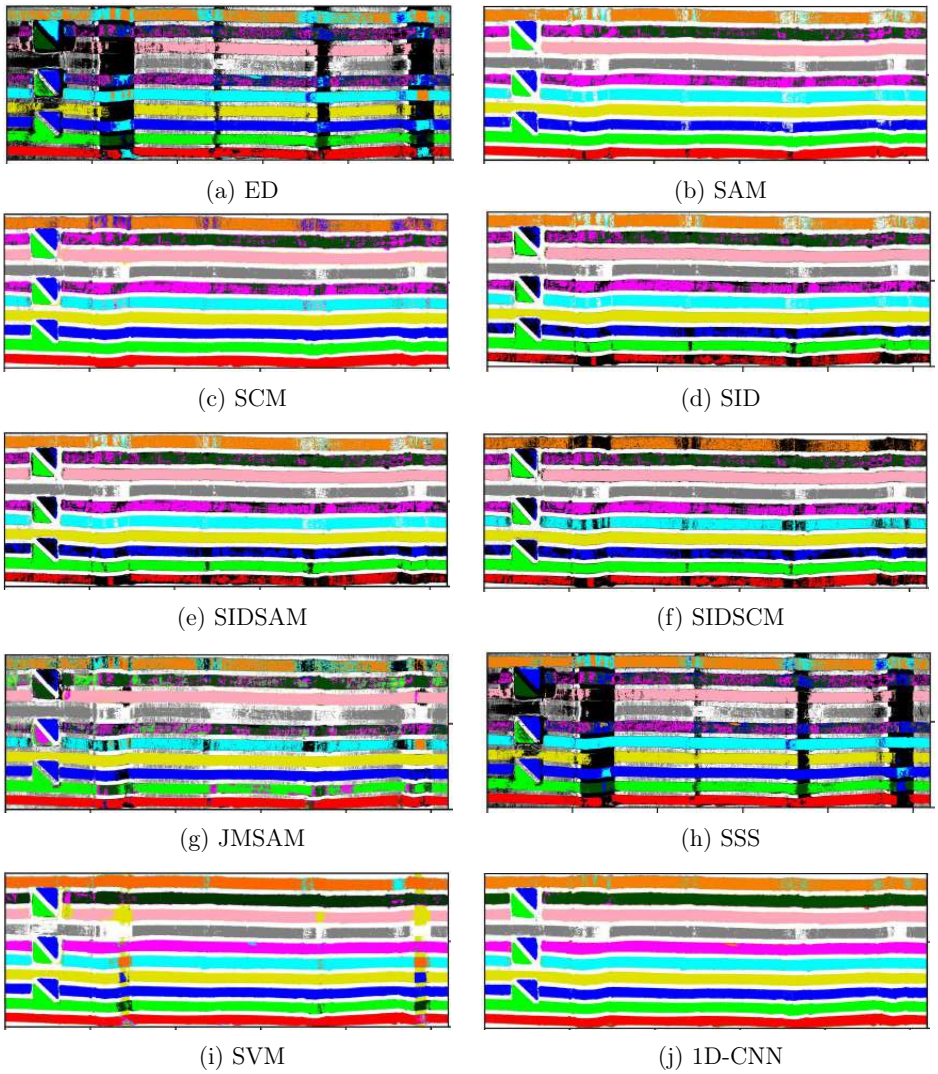
[74] Israel Cohen, Yiteng Huang, Jingdong Chen, Jacob Benesty, Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. Pearson correlation coefficient. *Noise reduction in speech processing*, pages 1–4, 2009.

[75] Luyan Ji, Peng Gong, Xiurui Geng, and Yongchao Zhao. Improving the accuracy of the water surface cover type in the 30 m from-glc product. *Remote Sensing*, 7(10):13507–13527, 2015.

[76] Augustine Gray and John Markel. A spectral-flatness measure for studying the autocorrelation method of linear prediction of speech analysis. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 22(3):207–217, 1974.

# Appendix

## A Classification results.



(a) ED (b) SAM

(c) SCM (d) SID

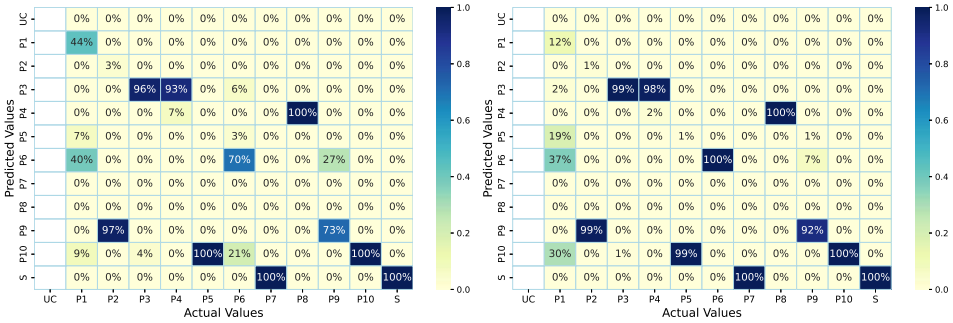
(e) SIDSAM (f) SIDSCM

(g) JMSAM (h) SSS

(i) SVM (j) 1D-CNN

Figure A.1: The classification results from ten algorithms using the reference spectrum from the flat region.

# B Confusion matrix for regions R3 and R13, illustrating the misclassification of pigments for SVM.




(a) (b)

Figure B.1: Confusion matrix for SVM; (a): ROI : 3 , (b): ROI : 13.

# C Confusion matrix illustrating classification accuracy obtained on a flat surface using SID Algorithm.

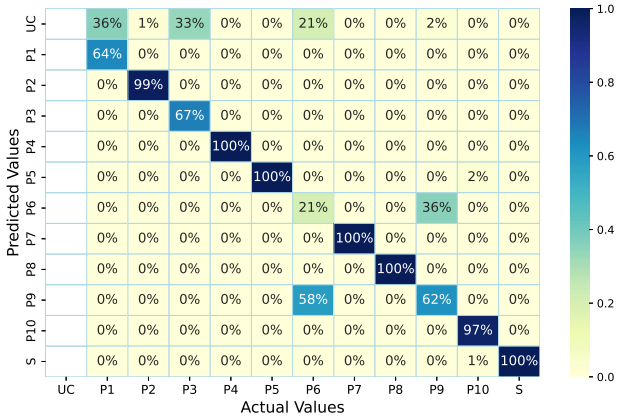
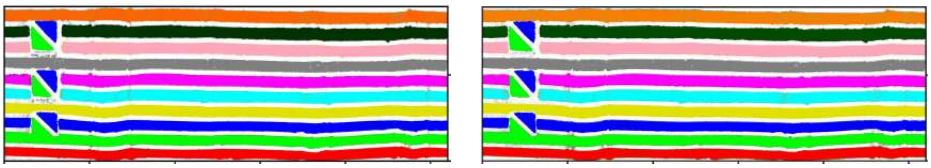



Figure C.1: Confusion matrix obtained by utilizing the SID Algorithm on a flat surface for ten pigments and substrate.

# D Classification results for SVM and 1D-CNN.



(a) SVM (b) 1D-CNN

Figure D.1: The classification results for SVM and 1D-CNN using the averaged reference spectrum.

# E Normalized reflectance spectrum of 10 pigments and substrate

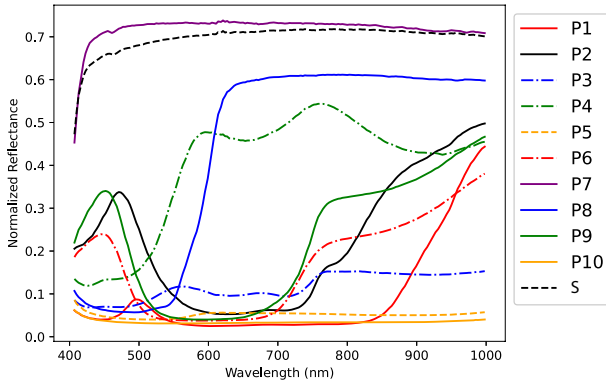



Figure E.1: Spectrum for ten pigments (P1 to P10) and substrate (S) measured at the flat surface by taking an average of 10 × 10 pixels.

# F Accuracy of Pigment Classification at Various Elevations Using Different Reference Spectra Conditions.

| | Flat Region | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| ED | 1 | 1 | 1 | 1 | 0.9 | 0 | 0.7 | 1 | 0.4 | 0.9 |
| | 78.74% | | | | | | | | | |
| SAM | 0.9 | 1 | 1 | 1 | 1 | 0.1 | 1 | 1 | 0.6 | 1 |
| | 85.34% | | | | | | | | | |
| SCM | 1 | 1 | 1 | 1 | 1 | 0.1 | 1 | 1 | 0.6 | 1 |
| | 86.25% | | | | | | | | | |
| SID | 0 | 1 | 0.8 | 1 | 1 | 0.1 | 1 | 1 | 0.6 | 1 |
| | 75.04% | | | | | | | | | |
| SID-SAM | 0 | 1 | 0.8 | 1 | 1 | 0.1 | 1 | 1 | 0.6 | 1 |
| | 74.60% | | | | | | | | | |
| SID-SCM | 0 | 1 | 0.8 | 1 | 1 | 0.1 | 1 | 1 | 0.6 | 0.9 |
| | 73.69% | | | | | | | | | |
| JMSAM | 1 | 0.9 | 0.9 | 1 | 1 | 0 | 0.6 | 1 | 0.3 | 0.9 |
| | 75.60% | | | | | | | | | |
| SSS | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0.4 | 1 |
| | 83.38% | | | | | | | | | |
| SVM | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 |
| | 99.30% | | | | | | | | | |
| 1D-CNN | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | 99.98% | | | | | | | | | |

| | Elevation : 2.5 mm | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| ED | 0.9 | 0.9 | 1 | 0.9 | 0.8 | 0.3 | 0.6 | 0.8 | 0.7 | 0.5 |
| | 74.82% | | | | | | | | | |
| SAM | 1 | 1 | 1 | 1 | 1 | 0.5 | 1 | 1 | 0.9 | 0.6 |
| | 88.31% | | | | | | | | | |
| SCM | 1 | 1 | 1 | 1 | 0.9 | 0.5 | 0.9 | 1 | 0.9 | 0.5 |
| | 87.38% | | | | | | | | | |
| SID | 0.2 | 1 | 0.9 | 1 | 1 | 0.5 | 1 | 1 | 0.9 | 0.7 |
| | 80.95% | | | | | | | | | |
| SID-SAM | 0.2 | 1 | 0.9 | 1 | 1 | 0.5 | 1 | 1 | 0.9 | 0.6 |
| | 80.11% | | | | | | | | | |
| SID-SCM | 0.3 | 1 | 0.9 | 1 | 0.9 | 0.5 | 0.9 | 1 | 0.9 | 0.4 |
| | 76.97% | | | | | | | | | |
| JMSAM | 0.8 | 0.7 | 1 | 1 | 0.7 | 0.2 | 0.9 | 1 | 0.6 | 0.4 |
| | 73.79% | | | | | | | | | |
| SSS | 1 | 1 | 1 | 1 | 1 | 0.3 | 0.8 | 1 | 0.8 | 0.6 |
| | 84.17% | | | | | | | | | |
| SVM | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | 99.83% | | | | | | | | | |
| 1D-CNN | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | 99.55% | | | | | | | | | |

| | Elevation : 5 mm | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| ED | 0.7 | 0.6 | 0.9 | 0.4 | 0.7 | 0.3 | 0.1 | 0.5 | 0.6 | 0.3 |
| | 51.14% | | | | | | | | | |
| SAM | 1 | 1 | 1 | 1 | 0.9 | 0.4 | 0.6 | 1 | 0.9 | 0.4 |
| | 80.59% | | | | | | | | | |
| SCM | 1 | 1 | 1 | 1 | 0.7 | 0.3 | 0.6 | 1 | 0.9 | 0.3 |
| | 77.59% | | | | | | | | | |
| SID | 0.3 | 1 | 0.8 | 1 | 0.9 | 0.4 | 0.6 | 1 | 0.9 | 0.4 |
| | 72.99% | | | | | | | | | |
| SID-SAM | 0.3 | 1 | 0.7 | 1 | 0.9 | 0.4 | 0.6 | 1 | 0.9 | 0.4 |
| | 71.58% | | | | | | | | | |
| SID-SCM | 0.4 | 1 | 0.7 | 1 | 0.5 | 0.4 | 0.6 | 1 | 0.9 | 0.2 |
| | 66.13% | | | | | | | | | |
| JMSAM | 0.7 | 0.4 | 1 | 1 | 0.5 | 0.2 | 0.5 | 1 | 0.6 | 0.3 |
| | 62.62% | | | | | | | | | |
| SSS | 0.8 | 0.7 | 1 | 0.5 | 0.8 | 0.3 | 0.2 | 0.5 | 0.6 | 0.4 |
| | 58.99% | | | | | | | | | |
| SVM | 1 | 1 | 1 | 1 | 0.9 | 1 | 0.8 | 1 | 1 | 1 |
| | 96.98% | | | | | | | | | |
| 1D-CNN | 1 | 1 | 1 | 1 | 0.9 | 1 | 0.9 | 1 | 1 | 1 |
| | 97.74% | | | | | | | | | |

| | Elevation : 10 mm | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| ED | 0.6 | 0.5 | 0.6 | 0.4 | 0.5 | 0.1 | 0.2 | 0.3 | 0.4 | 0.6 |
| | 41.61% | | | | | | | | | |
| SAM | 1 | 1 | 1 | 1 | 1 | 0.6 | 0.4 | 1 | 0.7 | 0.6 |
| | 83.84% | | | | | | | | | |
| SCM | 1 | 1 | 1 | 1 | 0.9 | 0.6 | 0.4 | 1 | 0.8 | 0.5 |
| | 81.73% | | | | | | | | | |
| SID | 0.1 | 0.9 | 0.7 | 1 | 1 | 0.6 | 0.5 | 1 | 0.8 | 0.7 |
| | 72.47% | | | | | | | | | |
| SID-SAM | 0.1 | 0.8 | 0.7 | 1 | 1 | 0.6 | 0.5 | 1 | 0.8 | 0.7 |
| | 70.81% | | | | | | | | | |
| SID-SCM | 0.4 | 0.8 | 0.6 | 1 | 0.8 | 0.6 | 0.4 | 1 | 0.8 | 0.3 |
| | 67.81% | | | | | | | | | |
| JMSAM | 0.9 | 0.4 | 0.8 | 1 | 0.4 | 0.1 | 0.4 | 0.5 | 0.6 | 0.3 |
| | 54.31% | | | | | | | | | |
| SSS | 0.9 | 0.5 | 1 | 0.4 | 1 | 0.1 | 0.3 | 0.4 | 0.5 | 0.7 |
| | 56.91% | | | | | | | | | |
| SVM | 1 | 0.8 | 1 | 0.5 | 0.6 | 1 | 0.4 | 0.6 | 0.7 | 1 |
| | 76.35% | | | | | | | | | |
| 1D-CNN | 1 | 1 | 1 | 1 | 1 | 1 | 0.7 | 1 | 1 | 1 |
| | 97.22% | | | | | | | | | |

Figure F.1: The classification accuracy from various algorithms used for evaluating ten different pigments on a flat surface and three elevated surfaces; spectral library and training datasets built using reference spectrum from an elevated region of 2.5 mm; the color gradient utilized in this figure employs green to represent higher accuracy, red to indicate lower accuracy, and white represents an accuracy of 50%.

| | Flat Region | | | | | | | | | | Elevation : 2.5 mm | | | | | | | | | | Elevation : 5 mm | | | | | | | | | | Elevation : 10 mm | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| **ED** | 1 | 1 | 1 | 1 | 0 | 0.6 | 0 | 1 | 0.5 | 1 | 0.9 | 1 | 0.9 | 0.9 | 0.2 | 0.4 | 0.5 | 0.8 | 0.3 | 0.9 | 0.7 | 0.6 | 0.7 | 0.3 | 0.3 | 0.2 | 0.3 | 0.5 | 0.3 | 0.9 | 0.7 | 0.5 | 0.4 | 0.4 | 0.2 | 0.4 | 0.2 | 0.3 | 0.2 | 0.9 |
| | **70.58%** | | | | | | | | | | **69.83%** | | | | | | | | | | **47.81%** | | | | | | | | | | **41.56%** | | | | | | | | | |
| **SAM** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 0.8 | 1 | 1 | 1 | 1 | 0.9 | 1 | 0.6 | 1 | 1 | 0.7 | 1 | 1 | 1 | 1 | 1 | 1 | 0.5 | 1 | 1 | 0.9 |
| | **99.57%** | | | | | | | | | | **97.03%** | | | | | | | | | | **92.07%** | | | | | | | | | | **92.55%** | | | | | | | | | |
| **SCM** | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 0.3 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 0.7 | 1 | 1 | 1 | 1 | 0.9 | 1 | 0.7 | 1 | 1 | 0.5 | 1 | 1 | 1 | 1 | 0.9 | 1 | 0.5 | 1 | 1 | 0.8 |
| | **91.23%** | | | | | | | | | | **95.24%** | | | | | | | | | | **91.15%** | | | | | | | | | | **92.08%** | | | | | | | | | |
| **SID** | 0.3 | 0.6 | 0.8 | 1 | 1 | 0.9 | 1 | 1 | 0.9 | 0.9 | 0.5 | 0.8 | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 0.9 | 0.8 | 0.4 | 0.8 | 0.8 | 1 | 0.9 | 0.8 | 0.7 | 1 | 0.9 | 0.7 | 0.8 | 0.6 | 0.8 | 1 | 1 | 0.9 | 0.5 | 1 | 0.9 | 0.9 |
| | **84.55%** | | | | | | | | | | **88.43%** | | | | | | | | | | **80.41%** | | | | | | | | | | **82.14%** | | | | | | | | | |
| **SID-SAM** | 0.4 | 0.7 | 0.8 | 1 | 1 | 0.8 | 1 | 1 | 0.9 | 0.7 | 0.5 | 0.8 | 0.9 | 1 | 0.9 | 0.9 | 1 | 1 | 0.9 | 0.8 | 0.4 | 0.8 | 0.8 | 1 | 0.9 | 0.8 | 0.7 | 1 | 0.9 | 0.7 | 0.8 | 0.5 | 0.7 | 1 | 1 | 0.9 | 0.5 | 1 | 0.9 | 0.9 |
| | **82.64%** | | | | | | | | | | **87.82%** | | | | | | | | | | **79.36%** | | | | | | | | | | **81.51%** | | | | | | | | | |
| **SID-SCM** | 0.8 | 1 | 0.8 | 1 | 0.1 | 0.8 | 1 | 1 | 1 | 0.1 | 0.7 | 1 | 0.9 | 1 | 0.6 | 0.9 | 1 | 1 | 0.9 | 0.5 | 0.6 | 1 | 0.7 | 1 | 0.9 | 0.7 | 0.7 | 1 | 0.9 | 0.5 | 0.9 | 0.7 | 0.7 | 1 | 0.8 | 0.9 | 0.5 | 1 | 0.9 | 0.6 |
| | **74.13%** | | | | | | | | | | **84.29%** | | | | | | | | | | **80.09%** | | | | | | | | | | **80.36%** | | | | | | | | | |
| **JMSAM** | 1 | 0.9 | 0.9 | 1 | 0 | 0.9 | 0 | 1 | 0.8 | 0.1 | 0.9 | 0.8 | 1 | 1 | 0.3 | 0.5 | 0.6 | 1 | 0.4 | 0.4 | 0.7 | 0.4 | 1 | 1 | 0.4 | 0.7 | 0.6 | 1 | 0.4 | 0.5 | 1 | 0.4 | 0.7 | 1 | 0.2 | 0.7 | 0.4 | 0.7 | 0.5 | 0.3 |
| | **66.85%** | | | | | | | | | | **68.74%** | | | | | | | | | | **66.34%** | | | | | | | | | | **58.83%** | | | | | | | | | |
| **SSS** | 1 | 1 | 1 | 1 | 0 | 0.7 | 0 | 1 | 0.6 | 1 | 1 | 1 | 1 | 1 | 0.4 | 0.4 | 0.5 | 1 | 0.4 | 0.9 | 0.8 | 0.7 | 1 | 0.5 | 0.6 | 0.2 | 0.3 | 0.5 | 0.3 | 0.8 | 0.8 | 0.5 | 0.7 | 0.4 | 0.3 | 0.5 | 0.2 | 0.4 | 0.2 | 0.9 |
| | **73.03%** | | | | | | | | | | **75.66%** | | | | | | | | | | **58.84%** | | | | | | | | | | **49.49%** | | | | | | | | | |
| **SVM** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.8 | 0.8 | 1 | 0.7 | 1 | 1 | 1 |
| | **99.08%** | | | | | | | | | | **99.85%** | | | | | | | | | | **99.70%** | | | | | | | | | | **92.93%** | | | | | | | | | |
| **1D-CNN** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | **99.98%** | | | | | | | | | | **99.56%** | | | | | | | | | | **98.81%** | | | | | | | | | | **99.34%** | | | | | | | | | |

Figure F.2: The classification accuracy from various algorithms used for evaluating ten different pigments on a flat surface and three elevated surfaces; spectral library and training datasets built using reference spectrum from an elevated region of 5 mm; the color gradient utilized in this figure employs green to represent higher accuracy, red to indicate lower accuracy, and white represents an accuracy of 50%.

| | Flat Region | | | | | | | | | | Elevation : 2.5 mm | | | | | | | | | | Elevation : 5 mm | | | | | | | | | | Elevation : 10 mm | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| ED | 1 | 0.8 | 0.9 | 0.2 | 0.3 | 0.5 | 0 | 0 | 0.5 | 1 | 0.9 | 0.7 | 1 | 0.6 | 0.6 | 0.4 | 0.2 | 0.5 | 0.3 | 0.6 | 0.7 | 0.7 | 0.8 | 0.6 | 0.6 | 0.2 | 0.5 | 0.6 | 0.3 | 0.4 | 0.5 | 0.7 | 0.7 | 0.3 | 0.4 | 0.4 | 0.2 | 0.3 | 0.2 | 0.7 |
| | 52.54% | | | | | | | | | | 57.61% | | | | | | | | | | 53.10% | | | | | | | | | | 44.44% | | | | | | | | | |
| SAM | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 1 | 0.8 | 1 | 1 | 0.8 | 1 | 0.9 | 1 | 0.9 | 1 | 0.7 | 0.9 | 1 | 1 | 0.9 | 1 | 1 | 1 | 0.7 | 1 | 0.7 | 0.9 |
| | 99.08% | | | | | | | | | | 93.20% | | | | | | | | | | 90.04% | | | | | | | | | | 91.67% | | | | | | | | | |
| SCM | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.7 | 0.7 | 1 | 1 | 1 | 1 | 0.9 | 1 | 0.9 | 1 | 0.6 | 0.5 | 1 | 1 | 1 | 1 | 1 | 1 | 0.8 | 1 | 0.6 | 0.7 |
| | 97.52% | | | | | | | | | | 92.98% | | | | | | | | | | 89.11% | | | | | | | | | | 91.12% | | | | | | | | | |
| SID | 0.2 | 1 | 0.7 | 1 | 1 | 0.6 | 1 | 1 | 0.9 | 1 | 0.5 | 1 | 0.5 | 1 | 1 | 0.9 | 1 | 1 | 0.7 | 0.8 | 0.4 | 1 | 0.6 | 1 | 0.9 | 0.7 | 1 | 1 | 0.7 | 0.7 | 0.6 | 1 | 0.7 | 1 | 1 | 0.9 | 0.8 | 1 | 0.8 | 0.9 |
| | 83.67% | | | | | | | | | | 83.32% | | | | | | | | | | 79.95% | | | | | | | | | | 86.31% | | | | | | | | | |
| SID-SAM | 0.2 | 1 | 0.6 | 1 | 1 | 0.5 | 1 | 1 | 0.8 | 0.8 | 0.5 | 1 | 0.4 | 1 | 1 | 0.9 | 1 | 1 | 0.6 | 0.8 | 0.4 | 1 | 0.5 | 1 | 0.9 | 0.6 | 0.9 | 1 | 0.7 | 0.7 | 0.5 | 1 | 0.6 | 1 | 1 | 0.9 | 0.8 | 1 | 0.7 | 0.9 |
| | 78.88% | | | | | | | | | | 81.36% | | | | | | | | | | 78.34% | | | | | | | | | | 83.92% | | | | | | | | | |
| SID-SCM | 0.6 | 1 | 0.5 | 1 | 0.9 | 0.4 | 1 | 1 | 0.9 | 0.1 | 0.6 | 1 | 0.4 | 1 | 0.9 | 0.8 | 1 | 1 | 0.6 | 0.5 | 0.7 | 1 | 0.6 | 1 | 0.9 | 0.6 | 0.9 | 1 | 0.7 | 0.4 | 0.7 | 1 | 0.7 | 1 | 1 | 0.9 | 0.8 | 1 | 0.7 | 0.6 |
| | 74.12% | | | | | | | | | | 78.73% | | | | | | | | | | 77.98% | | | | | | | | | | 83.15% | | | | | | | | | |
| JMSAM | 1 | 0.9 | 1 | 1 | 0.7 | 0.9 | 0 | 1 | 0.7 | 0.2 | 1 | 0.8 | 0.9 | 1 | 0.7 | 0.5 | 0.4 | 1 | 0.3 | 0.3 | 0.8 | 0.9 | 0.8 | 1 | 0.6 | 0.4 | 0.6 | 1 | 0.3 | 0.3 | 0.9 | 0.9 | 1 | 1 | 0.5 | 0.6 | 0.5 | 0.9 | 0.3 | 0.2 |
| | 72.79% | | | | | | | | | | 68.69% | | | | | | | | | | 66.79% | | | | | | | | | | 68.29% | | | | | | | | | |
| SSS | 1 | 1 | 1 | 0.6 | 1 | 0.6 | 0 | 0.3 | 0.6 | 1 | 1 | 0.7 | 1 | 0.6 | 1 | 0.5 | 0.4 | 0.5 | 0.3 | 0.7 | 0.9 | 0.7 | 0.9 | 0.7 | 0.8 | 0.2 | 0.5 | 0.6 | 0.3 | 0.5 | 0.6 | 0.8 | 1 | 0.4 | 1 | 0.4 | 0.3 | 0.4 | 0.2 | 0.8 |
| | 70.44% | | | | | | | | | | 67.06% | | | | | | | | | | 62.01% | | | | | | | | | | 58.22% | | | | | | | | | |
| SVM | 1 | 1 | 1 | 1 | 1 | 1 | 0.6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.8 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 0.9 | 0.9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | 95.48% | | | | | | | | | | 97.26% | | | | | | | | | | 96.94% | | | | | | | | | | 99.26% | | | | | | | | | |
| 1D-CNN | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | 99.97% | | | | | | | | | | 99.40% | | | | | | | | | | 98.51% | | | | | | | | | | 99.52% | | | | | | | | | |

Figure F.3: The classification accuracy from various algorithms used for evaluating ten different pigments on a flat surface and three elevated surfaces; spectral library and training datasets built using reference spectrum from an elevated region of 10 mm; the color gradient utilized in this figure employs green to represent higher accuracy, red to indicate lower accuracy, and white represents an accuracy of 50%.

| | Flat Region | | | | | | | | | | Elevation : 2.5 mm | | | | | | | | | | Elevation : 5 mm | | | | | | | | | | Elevation : 10 mm | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | P9 | P10 |
| **ED** | 1 | 1 | 1 | 1 | 0.4 | 0.5 | 1 | 1 | 0.5 | 0.9 | 1 | 1 | 1 | 0.9 | 0.7 | 0.4 | 0.3 | 0.8 | 0.3 | 0.6 | 0.7 | 0.7 | 0.9 | 0.4 | 0.6 | 0.2 | 0.1 | 0.5 | 0.3 | 0.4 | 0.6 | 0.5 | 0.6 | 0.4 | 0.4 | 0.4 | 0.2 | 0.3 | 0.2 | 0.6 |
| | **82.31%** | | | | | | | | | | **70.02%** | | | | | | | | | | **47.58%** | | | | | | | | | | **41.12%** | | | | | | | | | |
| **SAM** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 0.9 | 1 | 0.7 | 1 | 1 | 0.8 | 1 | 1 | 1 | 1 | 1 | 1 | 0.5 | 1 | 1 | 0.9 |
| | **99.84%** | | | | | | | | | | **97.85%** | | | | | | | | | | **92.98%** | | | | | | | | | | **93.34%** | | | | | | | | | |
| **SCM** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.7 | 1 | 1 | 1 | 1 | 0.7 | 1 | 0.6 | 1 | 1 | 0.5 | 1 | 1 | 1 | 1 | 0.9 | 1 | 0.5 | 1 | 1 | 0.7 |
| | **99.37%** | | | | | | | | | | **95.82%** | | | | | | | | | | **88.27%** | | | | | | | | | | **90.90%** | | | | | | | | | |
| **SID** | 0.7 | 1 | 0.9 | 1 | 1 | 0.8 | 1 | 1 | 1 | 1 | 0.5 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 0.9 | 0.9 | 0.3 | 0.9 | 1 | 1 | 0.9 | 0.7 | 0.7 | 1 | 1 | 0.8 | 0.6 | 1 | 1 | 1 | 1 | 0.9 | 0.5 | 1 | 1 | 0.9 |
| | **94.54%** | | | | | | | | | | **92.33%** | | | | | | | | | | **83.26%** | | | | | | | | | | **87.63%** | | | | | | | | | |
| **SID-SAM** | 0.8 | 1 | 0.9 | 1 | 1 | 0.8 | 1 | 1 | 1 | 1 | 0.5 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 0.9 | 0.9 | 0.3 | 0.9 | 1 | 1 | 0.9 | 0.7 | 0.7 | 1 | 0.9 | 0.8 | 0.6 | 1 | 0.9 | 1 | 1 | 0.9 | 0.5 | 1 | 1 | 0.9 |
| | **94.01%** | | | | | | | | | | **92.04%** | | | | | | | | | | **82.87%** | | | | | | | | | | **87.38%** | | | | | | | | | |
| **SID-SCM** | 1 | 1 | 0.9 | 1 | 1 | 0.7 | 1 | 1 | 1 | 0.3 | 0.7 | 1 | 1 | 1 | 0.9 | 0.9 | 1 | 1 | 0.9 | 0.6 | 0.5 | 1 | 0.9 | 1 | 0.7 | 0.7 | 0.7 | 1 | 0.9 | 0.4 | 0.7 | 1 | 1 | 1 | 0.9 | 0.9 | 0.5 | 1 | 0.9 | 0.7 |
| | **89.32%** | | | | | | | | | | **89.12%** | | | | | | | | | | **78.63%** | | | | | | | | | | **85.76%** | | | | | | | | | |
| **JMSAM** | 1 | 0.9 | 1 | 1 | 0.6 | 0.9 | 1 | 1 | 0.6 | 0.5 | 1 | 0.8 | 1 | 1 | 0.6 | 0.6 | 0.5 | 1 | 0.3 | 0.4 | 0.7 | 0.5 | 1 | 1 | 0.5 | 0.7 | 0.5 | 1 | 0.3 | 0.4 | 0.9 | 0.7 | 1 | 1 | 0.5 | 0.7 | 0.3 | 0.9 | 0.4 | 0.3 |
| | **85.00%** | | | | | | | | | | **71.60%** | | | | | | | | | | **66.47%** | | | | | | | | | | **67.98%** | | | | | | | | | |
| **SSS** | 1 | 1 | 1 | 1 | 1 | 0.6 | 1 | 1 | 0.6 | 1 | 1 | 1 | 1 | 1 | 1 | 0.5 | 0.4 | 1 | 0.3 | 0.8 | 0.8 | 0.7 | 1 | 0.5 | 0.8 | 0.2 | 0.2 | 0.5 | 0.3 | 0.6 | 0.7 | 0.5 | 1 | 0.4 | 1 | 0.4 | 0.2 | 0.4 | 0.2 | 0.8 |
| | **91.47%** | | | | | | | | | | **78.45%** | | | | | | | | | | **57.31%** | | | | | | | | | | **55.87%** | | | | | | | | | |
| **SVM** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | **99.97%** | | | | | | | | | | **99.71%** | | | | | | | | | | **99.36%** | | | | | | | | | | **99.74%** | | | | | | | | | |
| **1D-CNN** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | **99.94%** | | | | | | | | | | **99.54%** | | | | | | | | | | **98.88%** | | | | | | | | | | **99.59%** | | | | | | | | | |

Figure F.4: The classification accuracy from various algorithms used for evaluating ten different pigments on a flat surface and three elevated surfaces; using average reference spectra for supervised and ML models; the color gradient utilized in this figure employs green to represent higher accuracy, red to indicate lower accuracy, and white represents an accuracy of 50%.

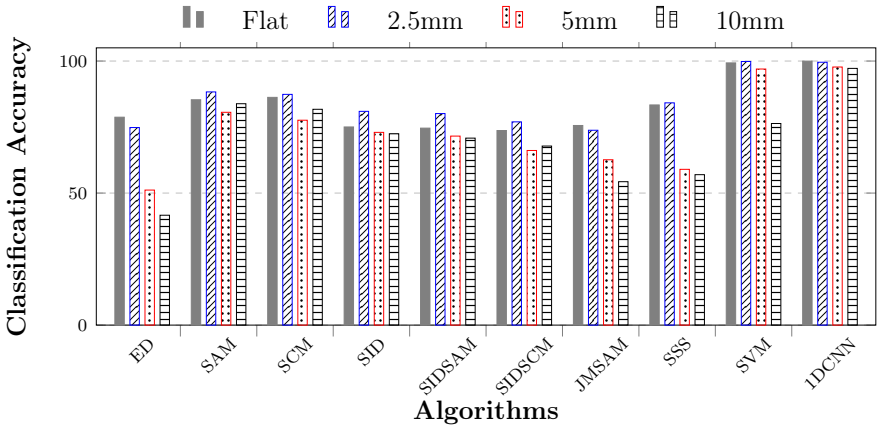



Figure F.5: Overall classification accuracy for four surface elevations (flat, 2.5mm, 5mm, and 10mm) obtained by using reference spectra taken from an elevated region of 2.5 mm.

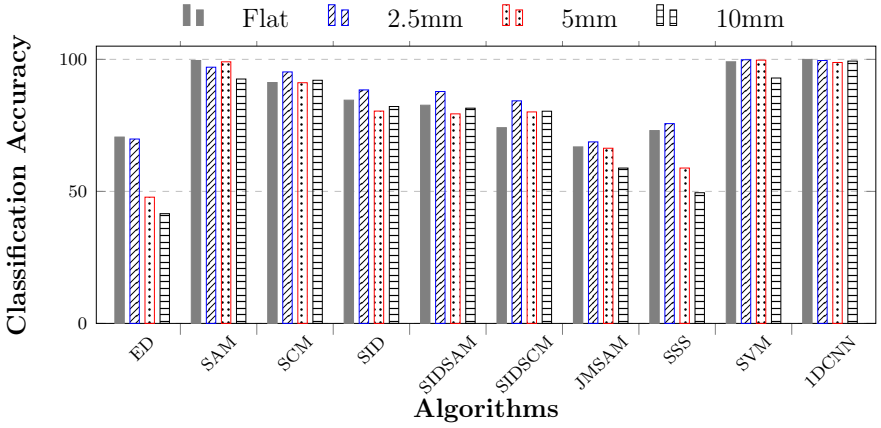



Figure F.6: Overall classification accuracy for four surface elevations (flat, 2.5mm, 5mm, and 10mm) obtained by using reference spectra taken from an elevated region of 5 mm.

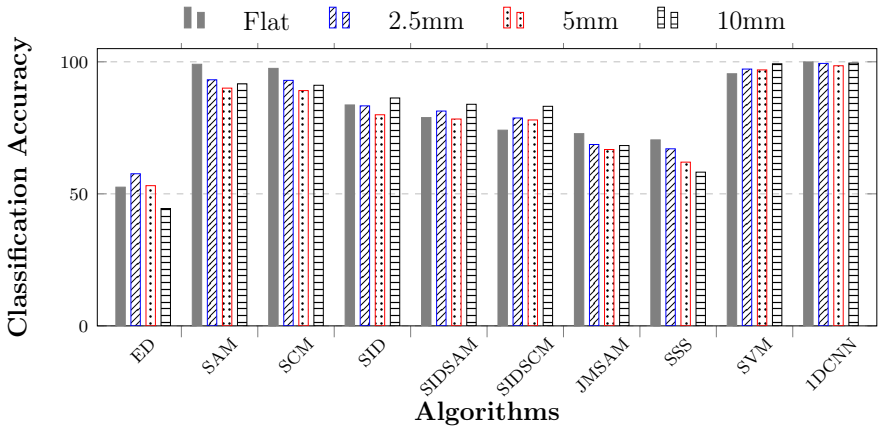



Figure F.7: Overall classification accuracy for four surface elevations (flat, 2.5mm, 5mm, and 10mm) obtained by using reference spectra taken from an elevated region of 10 mm.

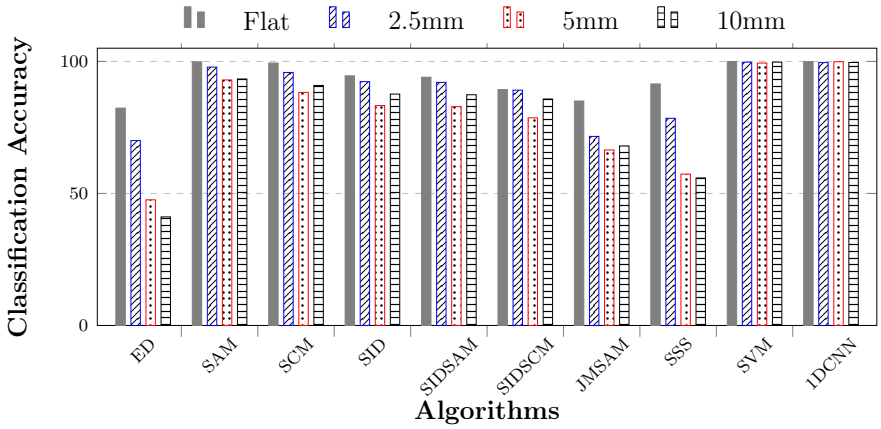
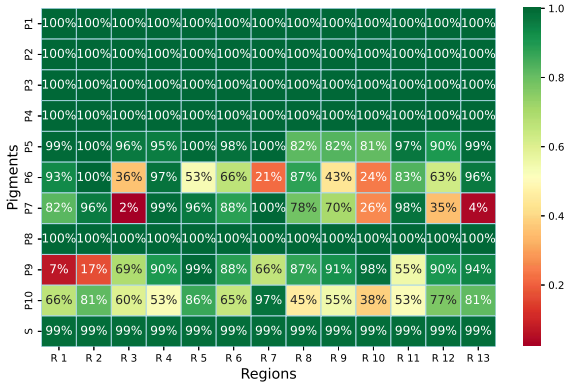
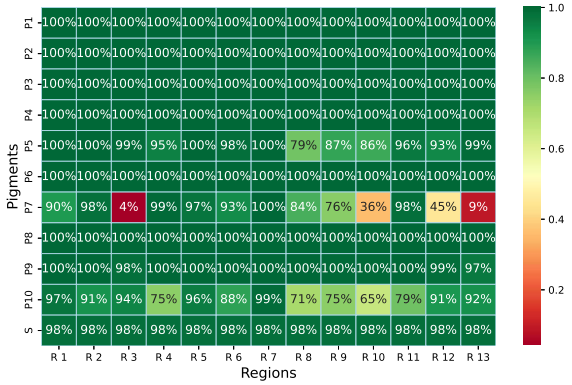



Figure F.8: Overall classification accuracy for four surface elevations (flat, 2.5mm, 5mm, and 10mm) obtained using average reference spectra; the averages were computed by considering a pixels values of a line drawn from right to left of the mockup for each pigment.

# G Classification Accuracy for SVM using Data Augmentation




(a)



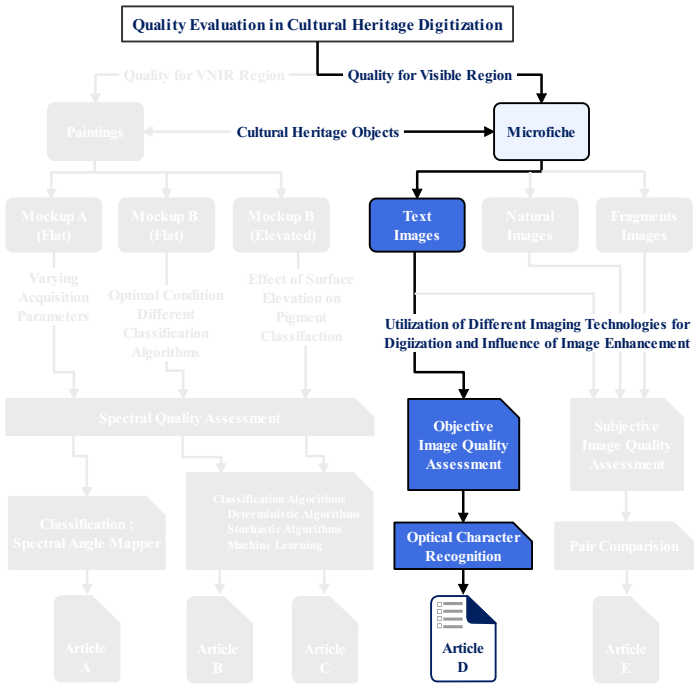

(b)

Figure G.1: Classification accuracy for each pigment across all Regions of Interest (ROIs) for SVM, (a): With data augmentation using single spectrum from flat region, and (b): With data augmentation using averaged single spectrum.

# Article D

## Evaluation of Text Legibility in Alternative Imaging Approaches to Microfiche Digitization

H. Deborah and D. J. Mandal (2021). “Evaluation of Text Legibility in Alternative Imaging Approaches to Microfiche Digitization.” In: *Proc. IS&T Archiving*, pp. 96–101. DOI: 10.2352/issn.2168-3204.2021.1.0.22



https://doi.org/10.2352/issn.2168-3204.2021.1.0.22


# Evaluation of Text Legibility in Alternative Imaging Approaches to Microfiche Digitization

***Hilda Deborah, Dipendra J. Mandal; Norwegian Colour and Visual Computing Laboratory, Department of Computer Science, NTNU–Norwegian University of Science and Technology; Gjøvik, Norway***

## Abstract

*Microfiche was a common format used in microforms reproductions of documents, extensively used for archival storage before the move to digital formats. While contemporary documents are still available for digitization, others from older historical periods are no longer physically accessible for various reasons. In some cases, their microfiche copies are available, making microfiche digitization a must. However, a microfiche reader is not always available and, even then, it is a machine made for the purpose of reading and not for data collection. In this work, the performance two imaging devices are evaluated as alternatives to the traditional microfiche reader, by means of optical character recognition (OCR). Results show that this alternative surpasses the performance of a microfiche reader in terms of text legibility.*

## Introduction

In the recent decades, we have seen an increase in the digitization of historical manuscripts using not only high-end color cameras and scanners, but also using multispectral [1], [2] and hyperspectral [3] imaging. There are significant advantages in doing so, not only for more accurate documentation purposes, but also for more advanced tasks, e.g., recovering hidden information [4]–[6]. Despite the need and advantages of such a digitization, especially for historical documents, there are cases where a *rescanning* of an object is no longer possible. Access to many historical documents in library collections across the globe can be difficult due to the fragile condition of the object. There are even cases where manuscripts or fragments have been lost [7]. Fortunately, when the documents have been kept in collections or institutions, often their records or analog copies are available in microforms. In this work, we focus on a specific format of microforms, i.e., microfiche.

Prior to the advance of digital technologies, microforms were the only available way to archive and preserve large documents. It was quickly adopted by the cultural heritage sector to capture their collection for preservation, access, and distribution. Microfiche is plasticky flat film sheets commonly used for reproducing historical printed documents, e.g., books and newspapers, in an optically reduced size or microforms [8]. These are of various types, e.g., silver-halide, diazo, and vesicular, and are available with different reduction ratios and life expectancy up to 500 years. The microfiche may be negative, i.e., clear lettering on a dark background, or its opposite, i.e., positive microfiche.

An example of a microfiche is provided in **Figure 1** and, taking note of its physical dimension, we can see that a single microfiche contains multiple photos or pages. Due to its significant reduction ratio, a microfiche reader [9], [10] is required to be able to observe and read its content or pages. Today's commonly available consumer or phone cameras would rarely have the resolution required to read a microfiche. This poses two challenges. The first is that a microfiche reader might not be as available as it was before since the technology has largely been replaced by digital technologies. A microfiche reader is also an analog machine made for the purpose of reading and not for digitization or data collection purposes. Thus, despite providing a high resolution, the use of a microfiche reader for digitizing microfiches is very time consuming. A single page in a microfiche equals a single image capture, requiring manual adjustments or placements of the lens such that it points to the right page. For one microfiche alone, use **Figure 1** as an example, 60 image captures are needed. And when talking about a digitization effort, we have hundreds if not thousands of microfiches, making the use of a microfiche reader impractical and highly costly. Additionally, it is also important to ensure the quality of the digitized images to meet user objectives.

The aim of this work is to find alternative imaging technologies for the digitization of microfiches. Trading off resolution with accessibility and time constraints, we are comparing two different imaging setups for the task of microfiche digitization. Information obtained from a microfiche must be readable. Thus, we define legibility as the evaluation criteria, and it is to be assessed by means of Optical Character Recognition (OCR) [11]–[13]. By using OCR, we limit the legibility assessment to system performance and, therefore, excluding assessment by human observers.

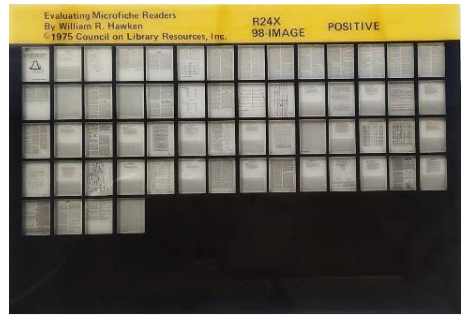



***Figure 1.*** *An example of a positive microfiche, of physical dimension 105mm × 148mm and a reduction ratio of 24X or 24 times, which will be used in the assessment of text legibility experiment. This material comes from Ref. [14].*

## Imaging Approaches

Access to a microfiche reader or any microform reader machine for that matter, is scarce even though the technology used to have a central role in archiving. The challenge of reading a microfiche is mainly related to its very reduced size. However, considering advances in optical devices in the past decade alone, it is highly likely that an alternative imaging solution is available. The first one to consider is a flatbed scanner, which nowadays are available with a resolution up to 6400 dpi. Then, if we also consider the availability of macro lenses coupled with a high-resolution camera, it might be sufficient to resolve the reduction ratio of a microfiche. Based on these considerations, we select a professional grade flatbed scanner and an in-house film scanning system that couples a monochrome camera and a macro lens as alternatives to a microform reader.

### *Microform reader*

A microform reader Zeutschel delta plus was available to us through the local library in Gjøvik, Norway. This device was marketed for public or professional use in digitizing all formats of microfilm and photographic materials. Microfiche is also listed as a compatible input type. This device is said to support a *reduction ratio* of 7X-105X. Reduction ratio expresses the linear relationship between the size of a document and its photographically reduced format or *microimage* [14]. For example, if a 10 cm object has been reduced 10X, it means the microimage is of size 1 cm. Other specifications of the reader machine that is relevant for this study can be seen in **Table 1**. Despite the high reduction ratio support, we consider the operation ease of using this device for digitization purposes to be low. This is mainly due to the need to manually position the lens for every single page within the microfiche, see maximum fiche per scan factor in the table, i.e., $1/n$ with $n = 98$.

### *Flatbed scanner*

A flatbed scanner used in this study is a professional grade scanner aimed for scanning films, i.e., Epson Perfection 4870 Photo. This scanner has up to 4800 dpi and provide an option to scan in transmissive or transparency mode, thus suitable for the purpose of microfiche scanning. The immediate advantage of its use is in time saving. Even though its throughput in **Table 1** is given in terms of seconds per line instead of per image, it still more efficient than a microfiche reader since it can scan two whole microfiches in one capture. This makes the operational ease high because of a significant reduce in time and efforts that are needed for the manual adjustments of apparatus and materials before each image capture.

### *In-house film scanner*

An in-house LED-based multispectral film scanner with the main purpose of capturing various kinds of film colors in transmission mode [15]. This scanner couples a monochrome camera with a macro lens, see details in **Table 1**, and therefore suitable for microfiche scanning. Since microfiche materials in this study is black and white, we only take grayscale images with one light source instead of multispectral images with the full range of the LED lights. The light source used was 415.5 nanometer, chosen arbitrarily but kept constant throughout the acquisition of all images. The maximum scan area of this scanner is not only due to the field of view of the scanner, but also due to how the apparatus is built for capturing images in transmissive mode. It has a square hole of roughly the size of a 35 mm film and only objects smaller than that size can be captured. For the specific test microfiche used in this study, the apparatus allows capturing six pages within a single microfiche. Nevertheless, both the throughput speed and operational ease can still be considered as high.

***Table 1.*** *Comparison of the specifications and characteristics of the three imaging setups evaluated in this study. Note that this summary is formulated within the specific context of reading 105 mm x 148 mm microfiches in a monochrome setup. The test microfiche has 24X reduction ratio and maximum $n = 98$ pages.*

| Factors | Microform reader | Flatbed scanner | In-house film scanner |
|---|---|---|---|
| **Model** | Zeutschel delta plus | Epson Perfection 4870 Photo | QHY600 16BIT BSI, atx-i 100mm F2.8 FF MACRO |
| **Compatible input types** | Microfiche, microcards, 16/35 mm roll microfilm, photographic slides, negatives, 35 mm perforated films | A4 size document, transparencies, photos, 35 mm films, negatives, 4”x5” formats | 35 mm photographs and motion picture films, small objects of different kinds |
| **Max. scan area** | 35 x 47 mm | 216 x 297 mm | 35 x 40 mm |
| **Max. fiche/ scan** | $1/n$ | $2n$ | $6/n$ |
| **Effective pixels** | 10 MP | 40,800 x 56,160 at 4800 dpi | 9,576 x 6,388 (±60 MP) |
| **Illumination** | Custom-calibrated LED array | Cold cathode fluorescent lamp | Calibrated LEDs |
| **Throughput speed** | Medium (±0.3 sec/ image) | High (±0.027 sec/ line) | High (±0.4 sec/ image) * |
| **Operation ease** | Low | High | High |

*The speed is calculated from specification of the camera given which was given as 2.5 fps for 16-bit output.

## Experimental Setup

The flowchart of assessing text legibility of the three imaging setups for the context of microfiche digitization in this study can be seen in **Figure 2**. Microfiche materials will be captured using the different devices, resulting in grayscale digital images. Note that despite the ability of these devices to capture color or multispectral images, it is unnecessary for the purpose of this experiment.

Depending on the experiment, a post-processing of the digital image might be carried out to remove noise and smooth an image by means of a median filter. Then, the images will be passed on to an optical character recognition (OCR) engine, which in this case is the open-source Tesseract-OCR. An OCR engine takes an image as input and return texts it can read from the input image. By comparing this recovered text with its corresponding ground truth, a text similarity measure using Levenshtein edit distance [16] will be calculated.

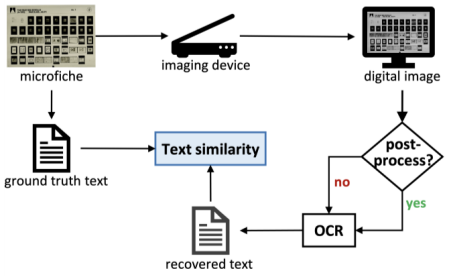
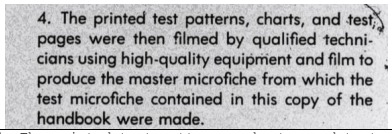



***Figure 2.*** *Experiment flowchart of the assessment of text legibility by means of Optical Character Recognition (OCR). Three imaging devices are compared, i.e., a microform reader, a flatbed scanner, and an in-house film scanner.*

### *Microfiche materials and their processing*

The test microfiche used in this experiment is one of the microfiches provided by a handbook for evaluating microfiche readers [14]. It is the positive microfiche with 24X size reduction and, therefore, allowing a single microfiche to contain a maximum of 98 images. The microfiche itself contains of only 60 images, see **Figure 1**, composed of the microimages of all pages in the handbook. Considering their relevance for text legibility assessment using OCR, only 16 pages are used. Fourteen pages used in the experiment are written in two-columns page. This poses a necessity to split the image of a page into its individual column to avoid confusion in the order of reading by OCR. Consequently, the ground truth text is also made following such order. A subset of a column and its corresponding ground truth text can be seen in **Figure 3**.


4. The printed test patterns, charts, and test pages were then filmed by qualified technicians using high-quality equipment and film to produce the master microfiche from which the test microfiche contained in this copy of the handbook were made.

```
4. The printed test patterns, charts, and test
pages were then filmed by qualified techni-
cians using high-quality equipment and film to
produce the master microfiche from which the
test microfiche contained in this copy of the
handbook were made.
```

***Figure 3.*** *A subset image of a microfiche-column and its ground truth text.*

### *Tesseract-OCR*

An optical character recognition (OCR) engine allows the conversion of digital images of typed, handwritten, or printed text into machine-encoded text, by recognizing a character at a time. It is particularly useful for automatizing a data entry process from printed records. In addition to working with images of documents, it can also be used to recognize text in a photograph of a scene. Tesseract-OCR is an open-source OCR engine that has been trained not only to detect single characters, but also optimized to recognize the shapes of letters for better recognition in case of blurred images. Furthermore, it also uses dictionary to improve text accuracy at the character segmentation step [17]. Our use of Tesseract-OCR is done through the Python wrapper `pytesseract`, and it returns the extracted text which will then be compared to the ground truth text.

### *Levenshtein edit distance*

We have seen an example of an image input for the OCR and its corresponding ground truth text in **Figure 3**. The accuracy of the text returned by the OCR, however, will vary depending on the quality of the input image. This further means that the accuracy depends on the quality of the imaging device. Using the last two lines from the image in **Figure 3** as an example, below are the texts returned by OCR for the exact image:

```
test microfiche contained in this oaPy of the
handbook were made. :
```

Comparing the above text to its ground truth in **Figure 3**, two mistakes can be spotted. The word `copy` is recognized as `oaPy` and there is also an extra colon (`:`). In computational linguistics, edit distance is used to quantify the difference between two texts by calculating the minimum number of *operations* required to transform one string to another. Different edit distances consider different operations in its calculation, e.g., deletion or substitution. Levenshtein edit distance (ED) [16] is chosen since it considers deletion, insertion, and substitution. This enables comparing two strings of different lengths unlike, e.g., Hamming distance [18]. Calculating the difference of the above text and its ground truth using ED, we obtain the score of 5. Since ED is a distance function, smaller value means higher text similarity, therefore indicating a better imaging setup for the task at hand. In addition to the standard ED, we will also use cumulative or aggregate ED to allow better comprehension of the overall performance of an imaging device.

## Results and Discussion

The results of legibility assessment of the three imaging setups for the use of microfiche digitization can be observed in **Figure 4**. In this graph, four entries are provided since two different dpi are evaluated for the flatbed scanner, i.e., at 4800 (Flatbed 4k) and 2400 (Flatbed 2k) dpi. By a quick observation, we can see that for certain microfiches, the use of a flatbed scanner with a 2400 dpi is insufficient to resolve the text from the microimage. Interestingly, it can also be seen that at 4800 dpi, the flatbed scanner almost always outperforms the microform reader. To have another point of view of

the performance, see **Figure 5**, where ED is plotted in a cumulative manner along the x-axis. A cumulative ED at any point of the x-axis is a sum of all EDs from the previous points. Here, it becomes clearer that the flatbed scanner at 4800 dpi outperforms the others and the in-house film scanner in general performs better than the microform reader.

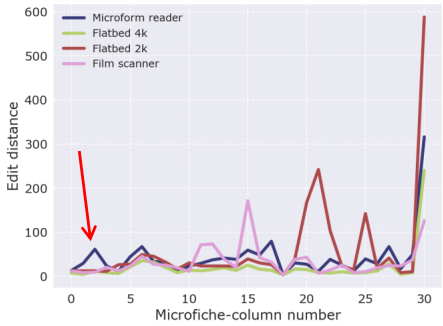



***Figure 4.*** *Levenshtein edit distance (ED) of the compared imaging setups, computed against each document's ground truth text. Two different dpi are evaluated for the flatbed scanner, i.e., Flatbed 4k and 2k.*

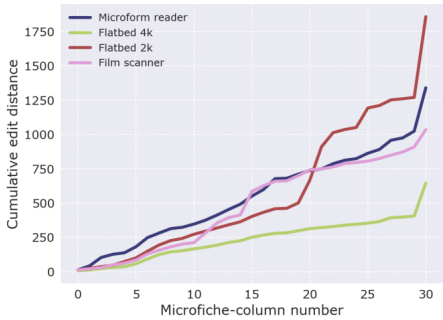



***Figure 5.*** *ED of the compared imaging setups, shown in a cumulative manner along the x-axis. It shows that the flatbed scanner at 4800 dpi is the best performing one at providing legible texts as evaluated by an OCR and that, interestingly, the microform reader is not clearly superior from the rest.*

To obtain a more thorough understanding of why and when a certain device is a better choice, an observation of the images is needed. Microfiche-column 3/5-1 is the one resulting in the first peak in the microform reader plot in **Figure 4**, as pointed by the red arrow. A subset area of that image can be observed in **Figure 6**. Upon a visual observation, both the contrast and sharpness of the flatbed scanner images in **Figure 6**(b)-(c) are significantly reduced compared to the one in **Figure 6**(a). The legibility score as measured by ED is, however, conversely related. Despite seemingly having a lower image quality, the flatbed scanner image at 4800 dpi has a lower ED of 11 compared to that of the microform reader with ED of 61. Even the image at 2400 dpi provides a better ED of only 12. The visual similarity of the flatbed scanner and film scanner images, those that provide low EDs, are the smoothness of the background. On the other, despite sharp letters, the microform reader image is noisy and granular in its background content. With the hypothesis that the background content tampers with the text recognition of the OCR, a smoothing filter can be used to improve the general performance of the microform reader and the film scanner.

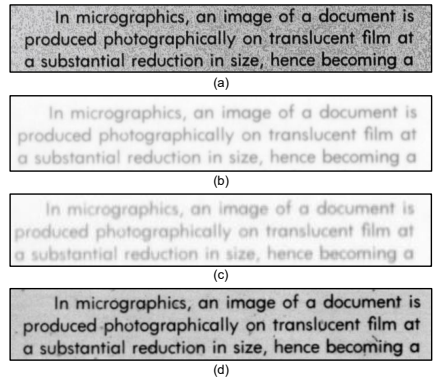



***Figure 6.*** *A subset area of microfiche-column 3/5-1 obtained from the compared (a) microfiche reader, (b) flatbed scanner at 4800 dpi and (c) 2400 dpi, and (d) in-house film scanner. The obtained ED scores for the microfiche-column for the respective devices are 61, 11, 12, and 9.*

After applying median filters of varying kernel size, the performance of the in-house film scanner is measured and plotted in **Figure 7**. We can see that the use of kernel size 3 improves the performance, albeit insignificantly. However, with kernel size 5, OCR struggles at recognizing the text in the images. The impact of smoothing on the images can also be observed through examples in **Figure 8**. Here, it becomes clear that smoothing with kernel size 5 blurs the individual letters unlike in size 3 where they are still sharp. Smoothing is also applied to the images from the microform reader and the performance can be observed in **Figure 9**. Note that here we choose to visualize it in terms of cumulative ED for ease of reading and clarity purposes. In the figure, it can be observed that smoothing increases the legibility of the text, although only up to kernel size 11. When reaching size 13, the legibility performance starts to decrease again as shown by MF-13 in the plot. The impact of median filters at these cutoff sizes to the images are shown in an example in **Figure 10**. Compared to the original image in **Figure 6**(a), both median filtered images show less granular artefacts in the

background, allowing improvements in the legibility aspect. Nevertheless, the improvement is only possible when the text itself is not blurred, which is the difference that can be observed between the two images in **Figure 10**.

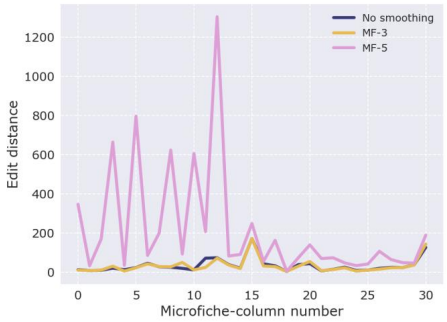



***Figure 7.*** *ED obtained by the in-house film scanner, at varying level of smoothing using median filters. MF-x in the figure legend means a median filter of kernel size x has been applied to the microfiche-column images.*

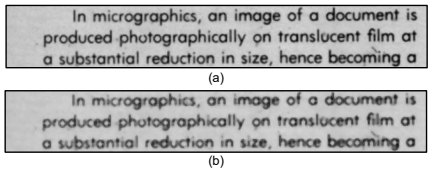



***Figure 8.*** *Subset area of microfiche-column 3/5-1 obtained from the in-house film scanner, and after applying a median filter of kernel size (a) 3 and (b) 5. The former gives the highest legibility, while the latter the lowest.*

Finally, taking the best combination of with or without smoothing, the performance of each imaging device in terms of text legibility can be observed through **Figure 11**. The flatbed scanner at 4800 dpi is still the best performing one. If we recall the initial result without smoothing as a post-processing step shown in **Figure 4** and **Figure 5**, we can now see that the performance of the microform reader has significantly improved when combined with a median filter of kernel size 11. Its result approximates the flatbed scanner at 4800 dpi. The improvement made by applying a median filter of size 3 to the in-house film scanner images, however, is insignificant. This makes its performance comes at the third place, after the flatbed scanner at 4800 dpi and the microform reader. Nevertheless, it is still a better choice for when the available flatbed scanner only provides a resolution of up to 2400 dpi.

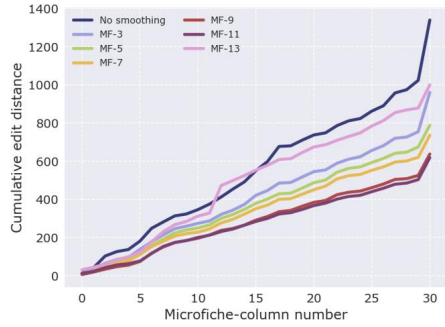



***Figure 9.*** *Cumulative ED obtained by the microform reader, after applying a median filter of kernel size x. The best performance is provided by applying a median filter of size 11, as shown by MF-11.*

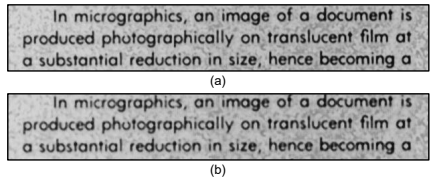



***Figure 10.*** *Subset area of microfiche-column 3/5-1 obtained from the microform reader, and after applying a median filter of kernel size (a) 11 and (b) 13. The former gives the highest legibility, while with the latter the legibility performance starts to decrease.*

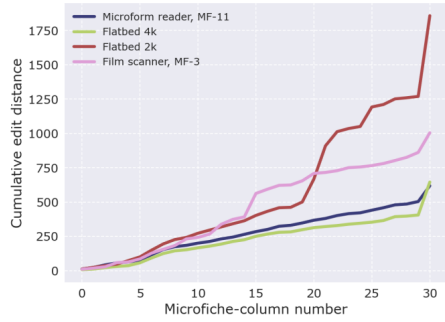



***Figure 11.*** *ED of the compared imaging setups, shown in a cumulative manner. Both microform reader and film scanner are combined with smoothing by means of median filters as a post-processing step.*

## Conclusion

In this study we have proposed a criterion for evaluating the quality of imaging devices for the task of microfiche digitization, i.e., text legibility. As an evaluation protocol, we have also proposed the use of OCR for an automatic recognition of the text in the images and Levenshtein edit distance as the metric. Three imaging devices have been compared, i.e., a microform reader, a flatbed scanner, and an in-house film scanner coupling a monochrome camera and a macro lens. As a conclusion, the flatbed scanner with 4800 dpi has been found to be the most suitable imaging device providing the highest quality of computer-legible texts.

This work has been motivated by our own research activities in the cultural heritage domain where, often, we do not have access to the physical objects for their rescanning using advanced imaging technologies. While in this study we have only assessed the legibility criteria, microfiche materials in our research are not only composed of written texts. There are also photographs and handwritten texts that will be unrecognizable by an off-the-shelf OCR. As a future work, we will develop more complete assessment protocols, considering other objective quality aspects as well as incorporating subjective evaluations by human observers.

## Acknowledgment

Authors would like to thank Torleif Elgvin, Gregory High, and Gjøvik Biblioteket for providing access to the various materials and devices used in the formulation and execution of this study. This work is supported by The Lying Pen of Scribes–Manuscript Forgeries, Digital Imaging, and Critical Provenance Research funded by the Research Council of Norway (projectnr. 275293) and CHANGE-ITN project funded by EU Horizon 2020 (Marie Skłodowska-Curie, grant agreement No. 813789).

## References

[1] R. L. (Jr. ) Easton and W. Noël, "The Multispectral Imaging of the Archimedes Palimpsest," *Gaz. Livre Médiév.*, vol. 45, no. 1, pp. 39–49, 2004, doi: 10.3406/galim.2004.1646.

[2] B. Stegmann, "Collaborative Manuscript Production and The Case of Reykjabók: Paleographical and Multispectral Analysis," in *New Studies in the Manuscript Tradition of Njáls saga*, Medieval Institute Publications, 2018, pp. 29–54.

[3] I. M. Cortea, L. Ghervase, L. Ratoiu, and R. Rădvan, "Application of Spectroscopic and Hyperspectral Imaging Techniques for Rapid and Nondestructive Investigation of Jewish Ritual Parchment," *Front. Mater.*, vol. 7, p. 601339, 2020, doi: 10.3389/fmats.2020.601339.

[4] A. Tournié *et al.*, "Ancient Greek text concealed on the back of unrolled papyrus revealed through shortwave-infrared hyperspectral imaging," *Sci. Adv.*, vol. 5, no. 10, 2019.

[5] L. Snijders, T. Zaman, and D. Howell, "Using Hyperspectral Imaging to Reveal a Hidden Precolonial Mesoamerican Codex," *J. Archaeol. Sci. Rep.*, vol. 9, pp. 143–149, 2016.

[6] A. Tonazzini *et al.*, "Analytical and mathematical methods for revealing hidden details in ancient manuscripts and paintings: A review," *J. Adv. Res.*, vol. 17, pp. 31–42, 2019.

[7] E. Tigchelaar, "On the Unidentified Fragments of 'DJD' XXXIII and PAM 43.680: A New Manuscript of '4QNarrative and Poetic Composition', and Fragments of '4Q13', '4Q269', '4Q525' and '4QSb'(?)," *Rev. Qumrân*, vol. 21, no. 3 (83), pp. 477–485, 2004.

[8] W. De Haas, "The Microfiche," *Am. Doc.*, vol. 9, no. 2, pp. 99–106, 1958, doi: 10.1002/asi.5090090204.

[9] C. Frost, "Microfiche readers," *Computer*, vol. 5, no. 3, pp. 58–60, 1971, doi: 10.1109/C-M.1972.216919.

[10] F. L. Keeler and W. A. Rizzo, "An evaluation of microfiche reader types for use with programmed instruction," Naval Training Equipment Center, TAEG-R-35, 1976.

[11] H. Deborah, H. M. Manurung, and A. M. Arymurthy, "Objective Criteria for Typewritten Old Document Image Enhancement and Restoration," *Proceeding Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACSIS*, 2010.

[12] J. Owen, "Using Optical Character Recognition to Identify Legibility of Non-western Languages," California Polytechnic State University, 2011.

[13] M. R. Gupta, N. P. Jacobson, and E. K. Garcia, "OCR binarization and image pre-processing for searching historical documents," *Pattern Recognit.*, vol. 40, no. 2, pp. 389–397, 2007, doi: 10.1016/j.patcog.2006.04.043.

[14] W. R. Hawken, *Evaluating Microfiche Readers: A Handbook for Librarians*. Council on Library Resources, 1975.

[15] G. Trumpy, S. George, J. Y. Hardeberg, and B. Flueckiger, "Multispectral capture of film colors with LEDs," in *Colour Photography and Film: Sharing knowledge of analysis, preservation, conservation, migration of analogue and digital materials*, 2021, p. 28.

[16] V. I. Levenshtein, "Binary Codes Capable of Correcting Deletions, Insertions and Reversals," *Sov. Phys. Dokl.*, vol. 10, p. 707, 1966.

[17] R. Smith, "An Overview of the Tesseract OCR Engine," in *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, Sep. 2007, vol. 2, pp. 629–633.

[18] R. W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Tech. J.*, vol. 29, no. 2, pp. 147–160, 1950.

## Author Biography

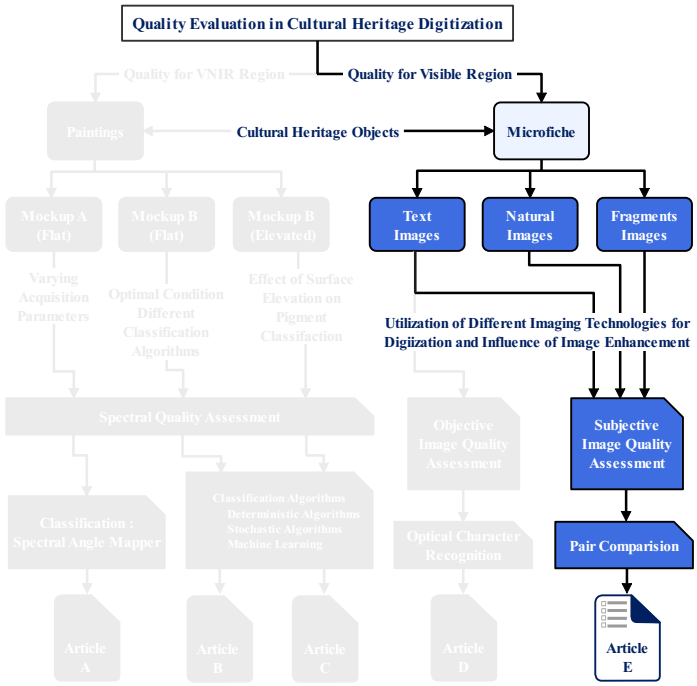
*Hilda Deborah received her BSc from the University of Indonesia (2010), MSc from the Erasmus Mundus Color in Informatics and Media Technology (2013), and PhD from NTNU and University of Poitiers (2016). She was a Marie Curie Postdoctoral Fellow at the University of Iceland and NTNU from 2018-2020. She is currently Senior Researcher at NTNU, with interests in hyperspectral imaging and fundamental image processing, with applications in the cultural heritage and remote sensing domains.*

*Dipendra Jee Mandal received his BE (2010) and MS (2017) degree in Electrical and Electronics Engineering from Kathmandu University, Nepal. He was an exchange student under Erasmus Mundus at University Lumiere Lyon 2, France (2017-2018). He is currently a Ph.D. student at NTNU under Marie Curie fellowship (CHANGE-ITN), his research field is quality assessment for cultural heritage digitization using hyperspectral imaging.*

# Article E

## Subjective Quality Evaluation of Alternative Imaging Techniques for Microfiche Digitization

D. J. Mandal, H. Deborah, and M. Pedersen (2023). “Subjective Quality Evaluation of Alternative Imaging Techniques for Microfiche Digitization.” In: *Journal of Cultural Heritage*. Accepted for publication



# Subjective Quality Evaluation of Alternative Imaging Techniques for Microfiche Digitization

Dipendra J. Mandal[1], Hilda Deborah[1], and Marius Pedersen[1]

[1]Department of Computer Science, Norwegian University of Science and Technology (NTNU), Norway

**Abstract**

Before the advent of digital formats, microfiche, a type of microform, was widely utilized for archiving and preserving historical documents. As a result, numerous historical collections and documents can only be found in microfiche format, transforming them into valuable artifacts and indispensable aspects of our cultural heritage. Although microfiche can last a long time, it is still susceptible to damage and requires digitization for preservation and broader accessibility. In addition, traditional microfiche readers are not always available and are primarily designed for reading rather than digitization. In this study, we evaluated the performance of two alternative imaging devices compared to a traditional microfiche reader and the impact of enhancement on image quality using subjective image quality assessment. The experiments were carried out in a controlled environment with twenty-one participants, including an expert. Our results showed that the reproduction of alternative devices was preferred over that of a traditional microfiche reader. Furthermore, our results demonstrate that image enhancement techniques significantly improved image quality. This study suggests that alternative imaging devices may be a viable option for digitizing microfiche and improving access to historical collections.

Keywords – *Microfiche, Image Quality Assessment, Subjective experiment, Dead Sea Scrolls*

## 1 Introduction

Microform refers to a method of storing information, such as a document or image, in a much smaller form than the original. Information can be stored on a microfilm, microfiche, or other types of microforms. This storage method is useful because it saves space and makes it easier to store large amounts of information in a compact and manageable form. Additionally, the microform can be easily reproduced, allowing for the creation of multiple copies that multiple users can access. Microfiche is a type of microform that consists of a sheet of film containing multiple microimages in a reduced form (commonly referred to as the reduction ratio) and is arranged in a

grid-like format [1]. The reduction ratio refers to the extent to which a document is visually scaled down (miniaturized) through photography. It is represented as a ratio between the original linear size and the linear size of the microform image. For instance, if an item is filmed at a reduction ratio of 20 to 1 (20:1), it means it has been reduced in size by a factor of twenty. Reduction ratios are often denoted as 20x, 40x, and so forth. Before the advent of advanced digital technologies, microforms were the exclusive method for archiving and conserving extensive documents, such as newspapers. The cultural heritage domain quickly adopted this technique to capture its collections to ensure preservation, ease of access, and greater distribution. Today, many historical collections can be found only on microforms. For example, Papua New Guinea's colonial-era history is only available in the form of microfiche [2].

Microforms contain an emulsion layer (a thin coating of light-sensitive material) embedded in a base: cellulose (typically nitrate, acetate, or triacetate) or polyester (plastic). Because of its flammability and rapid deterioration, cellulose nitrate was replaced by safer alternatives such as cellulose acetate and cellulose triacetate. The acetate film is prone to rapid deterioration due to fluctuations in temperature and humidity, causing distortion and warping of the emulsion. This deterioration produces acetic acid, resulting in the vinegar syndrome [3]. With better stability, durability, and tear resistance, polyester film became the preferred choice for modern microform production in the 1980s [4, 5]. Three main types of emulsion layers used in microforms are silver halide, diazo, and vesicular [4]. Silver halide films are highly light-sensitive, capturing extensive detail and offering a wide tonal range. When properly prepared and stored, they can last up to 500 years [6]. However, diazo films are prone to image fading and loss, with a useful life of approximately 100 years, and are not considered archival quality [7]. Vesicular films offer scratch resistance but can distort under intense use or high heat, lasting between 10 to 100 years when stored appropriately [8].

Microfiche materials, despite their longer life span, are prone to physical degradation caused by various factors [9]. Exposure to light, fluctuating temperatures and humidity levels, improper handling, and poor storage conditions can lead to the deterioration of the microfiche. This degradation often manifests as brittleness in the film base, resulting in compromised image quality or even complete disintegration. Furthermore, microfiche collections are vulnerable to environmental hazards such as water damage, mold growth, pests, and fire. These hazards can lead to the loss of information and render microfiche unusable. Thus, digitizing microfiche can be a valuable solution for addressing such preservation challenges. Nevertheless, preserving the original microfiche could still be necessary for specific purposes. Therefore, a combined approach to digitization and proper storage measures is essential.

Gaining access to numerous historical documents stored in libraries worldwide can prove challenging due to the fragile state of the objects. In some instances, manuscripts or fragments have even been lost [10]. This means that their microfiche copies might be the last surviving records of the objects or the only accessible option for the wider public. However, microfiche cannot be read directly by the human eye and requires a special device for enlarging, printing, and scanning the microforms into a readable format. Only a few specialized archives or libraries still possess microfiche readers to facilitate access to their microform collections.

With the adoption of digital technologies, the digitization of microfiche can offer numerous benefits [11]. With digital formats, they can be accessed and viewed on computers and other digital devices, making it easier to share microfiche collections with a wider audience, regardless of their location. Digitization also makes storing and managing microfiche collections easier, since they no longer require a physical storage space. Additionally, digitization can help to preserve the information in microfiche documents by reducing wear and tear on the physical copies and by making backup copies to guard against loss or damage. Digitization often results in reproductions that differ from the original in various ways. Digitizing microfiche, which requires magnification (the usual reduction factors are from 24, 42 or 48, 96 to 1 [12]), is likely to introduce attributes that can affect the final quality of the reproduction, e.g., noise, distortion, and artifacts. Figure S1, included as a supplementary file, provides a visual illustration of such artifacts showing an original document image and a magnified image obtained from the microfiche. Quality assessment should be included as an essential part of the digitization process to ensure that the digitized object maintains completeness, fidelity, and legibility compared to the original. Both objective and subjective methods [13, 14] can be used to assess image quality.

So far, to our knowledge, very few studies have been conducted in the quality assessment of microfiche digitization. Some of the work we found on microfiche quality assessment have concentrated on comparing the quality of the original paper document with its microfiche version. One such study, conducted by Lee et al. [15], used various quantitative measures to evaluate the fidelity of the original paper document and its 35 mm microfilm copies. The study results indicated that the copy was not fully faithful to the original document. In another study by Duff et al. [16] on the early Canadiana material collection, a comparison was made between the original materials and their microfiche counterparts. The findings showed a significant difference between the two formats, with most of the participants preferring the paper format over the digital version. As of today, the original of these are not easily accessible and, in some cases, may not even be available.

Numerous initiatives to convert microfiche into digital format are underway worldwide, utilizing established and innovative digitization methods and technologies. Scanning documents as images and then utilizing Optical Character Recognition (OCR) algorithms to convert the images into text is a widely employed approach for digitizing historical text collections [17, 18]. Recently, in the year 2020, the National Security Research Center (NSRC) conducted a trial of an artificial intelligence/ machine learning system aimed at digitizing certain microfilm and microfiche document collection [19]. Deborah and Mandal [20] used objective quality assessment to evaluate the performance of two imaging devices as alternatives to the traditional microfiche reader. They assessed the legibility of the images obtained by these devices when inputted to an optical character recognition (OCR) and used Levenshtein distance [21] as the text similarity measure. The results showed the superiority of a flatbed scanner at 4k dpi resolution to a traditional microform reader. It is also worth mentioning that to achieve those results, the images obtained by the microform reader had to undergo a post-processing step, meanwhile no enhancement was applied to those from the flatbed scanner. Since the study utilized an OCR, the focus was on microfiche containing typewritten text. However, microfiche materials may also contain handwritten texts that off-the-shelf OCR systems cannot recognize and also photographs where the legibility evaluation may be less relevant. To address this limitation, our study incorporated microfiche materials containing typewritten text, photographs of natural scenes, and photographs of ancient handwritten fragments. Moreover, we conduct a subjective quality assessment with human observers to evaluate the quality of the scanned images.

Enhancing images through image processing is a widely used practice that aims to improve their visual quality. By adjusting image attributes, enhancement can produce a more aesthetically pleasing result for a given scenario [22]. The contrast of an image is widely recognized as a significant quality attribute [23]. Improving contrast is commonly believed to enhance the perceived quality of most natural images [24, 25, 26]. Contrast Stretching (CS) is a technique commonly used to enhance low-contrast images, which involves using a piecewise linear curve to expand the dynamic range of gray levels [27]. In our experiment, we applied CS as a post-processing to find out if it enhances the quality of microfiche materials.

## 2 Research Aim

Many historical collections and documents can only be found in microfiche format, making them valuable artifacts of cultural heritage. Microfiche requires microfiche

readers, an imaging device that can enlarge the content of microfiche, allowing it to be read directly by the human eye. Such devices are primarily designed for reading rather than digitization, and only a limited number of specialized archives or libraries may still have them available. In addition, over time, these devices can suffer from various problems that affect their usability. Some common issues include yellow screens (due to aging and prolonged exposure to light), defective reels, degraded image quality, and mechanical failures. These problems can significantly impact the ability to access and read microfiche materials effectively. For instance, a discolored screen can negatively impact the contrast and readability of the microfiche images. Given the importance of preserving and accessing microfiche materials, exploring alternative options that can enhance the reading experience and facilitate the efficient utilization of microfiche becomes crucial. Therefore, this research aims to identify alternative imaging devices compared with traditional microfiche readers as a viable option for digitizing microfiche and improving access to historical collections. To incorporate a wider variability in the types of records, we used microfiche that contains typewritten text, photographs of natural scenes, and photographs of ancient handwritten fragments.

# 3 Materials and Methods

## 3.1 Imaging Approaches

Access to microfiche readers or other similar microform reader machines is limited today, although they were once a central aspect of archiving. The main challenge with reading microfiche is its significantly reduced size. However, with advances in optical technology in recent years, it is highly probable that alternative imaging solutions are available. The first option to consider is a flatbed scanner, which now offers resolutions up to 6400 dpi. Additionally, by using macro lenses alongside a high-resolution camera, it may be possible to resolve the reduction ratio of microfiche. Based on these considerations, we have opted to utilize a professional grade flatbed scanner at 4800 dpi (FBS) and an in-house film scanning (IFS) system, which includes a monochrome camera and a macro lens, as viable alternatives to a microform reader (FSL). Table 1 summarizes the specifications and characteristics of the imaging devices used in the experiment. For further details, we refer the reader to Deborah and Mandal [20].

Table 1: Specification and characteristics comparison of three imaging setups. Note that this summary is formulated within the specific context of reading 105 mm × 148 mm microfiches in a monochrome setup.

| **Factors** | **Microform reader** | **Flatbed scanner** | **In-house film scanner [28]** |
|---|---|---|---|
| **Model** | Zeutschel delta plus | Epson Perfection 4870 Photo | QHY600 16BIT BSI, atx-i 100mm F2.8 FF MACRO |
| **Compatible input types (non-exhaustive)** | Microfiche, microcards, 16/35 mm roll microfilm, photographic slides, negatives, 35 mm perforated films | A4 size document, transparencies, photos, 35 mm films, negatives, 4"×5" formats | 35 mm photographs and motion picture films, small objects of different kinds |
| **Max. scan area** | 35×47 mm | 216×297 mm | 35×40 mm |
| **Max. fiche per scan** | $1/n$ | $2n$ | $6/n$ |
| **Effective pixels** | 10 MP | 40,800×56,160 at 4800 dpi | 9,576×6,388 (±60 MP) |
| **Illumination** | Custom calibrated LED array | Cold cathode fluorescent lamp | Calibrated LEDs |
| **Throughput speed** | Medium (±0.3 sec/ image) | High (±0.027 sec/ line) | High (±0.4 sec/ image) * |
| **Operation ease** | Low | High | High |

*The speed was determined based on the camera's provided specification of 2.5 frames per second (fps) for 16-bit output.

## 3.2 Microfiche Materials

The experiment used microfiche materials from three different sources. They were microfiches provided by 1) a handbook for evaluating microfiche readers [29], 2) the Allegro Qumran Collection on Microfiche [30], and 3) the Dead Sea Scrolls on Microfiche [31]. The microfiche from the first source primarily contained textual content, including typewritten text with various fonts and font sizes. On the other hand, the other two microfiche sources consisted of scroll fragments containing biblical and non-biblical writings. Additionally, the microfiche included photographs of Qumran caves and natural scenes that capture images of natural landscapes and outdoor environments. Figure 1 illustrates an example of the figure obtained from the IFS device for three distinct microfiches used in an experiment. A summary of the technical specifications for the microfiche is presented in Table 2. It is important to note that the performance of the devices could be impacted by settings.


(a) (b) (c)

Figure 1: Illustration of microfiches obtained using the IFS device. (a) Text from a Handbook for Evaluating Microfiche Readers, (b) a natural scene from the Allegro Qumran Collection, and (c) fragments from the Dead Sea Scrolls.

## 3.3 Image Acquisition

Each of the three microfiches (containing text, natural scenes, and fragments) was scanned and captured using all three devices. We used a microform reader (Zeutschel Delta Plus), made available through the local library of Gjøvik, Norway. This device required manual lens adjustment for each page within the microfiche. The resulting images were saved in TIF format. We used a professional-grade flatbed scanner manufactured by Epson, which allowed us to scan two complete microfiches in a single capture. This significantly improved operational efficiency by eliminating the

Table 2: Summary of microfiche specification used in this experiment. The assumption is that the accompanying microfiche corresponds to the same time period based on its publication date.

| **Microfiche** | **Type** | **Reduction Ratio** | **Sheet per Fiche** | **Time Period** |
|---|---|---|---|---|
| Handbook for Evaluating Microfiche Readers | positive microfiche | 1:20 | 60 | 1975 |
| Allegro Qumarn Collection | positive silver halide microfiche | 1:19 | 50 | 1996 |
| Dead Sea Scrolls | positive silver halide microfiche | 1:13 | 50 | 1992 |

need for manual adjustments of equipment and materials prior to each image capture. After scanning, each page of the microfiche was cropped and saved as a TIF file. The in-house multispectral film scanner used for this study is equipped with an LED-based system and a monochrome camera with a macro lens. Grayscale images were captured using a single light source at a wavelength of 415.5 nanometers. The maximum scanning area is determined by its field of view, and the design is optimized to capture images in a transmissive mode. It allowed capturing up to six pages within a single microfiche per scan. Refer to Figure S2, added as a supplementary file, for a detailed setup schematic.

## 3.4 Psychovisual experiment

A subjective experiment was designed to evaluate the overall image quality of the digitized microfiche using three different imaging devices (FSL, FBS, and IFS). Several psychometric methods exist in the literature for measuring image quality [32, 33, 34, 35], such as paired comparison, rank ordering, categorical sort, and magnitude estimation. In this study, a force-choice pair comparison [36] experiment was designed without ties, i.e., observers were forced to choose one of the two preferences randomly if they found a tie between the stimuli. This method was chosen because of its relative simplicity compared to other methods and because it is better at finding differences between images. In a pair comparison experiment, the task is typically to indicate the preferred option from each pair of stimuli rather than assigning a quality score to each stimulus. A web-based tool called QuickEval [37] was utilized to carry out the experiment. This tool is specifically designed to carry out psychometric scaling experiments related to image quality.

### 3.4.1 Observers

In this study, there were a total of 21 observers, including one expert. The expert observer had experience with imaging, old/historical manuscripts, and microfiche, all of which were used in the experiment. Among the remaining observers, five had experience with old or historical manuscripts, and one had worked with microfiche. Moreover, most of these observers had experience in imaging. For the study, all observers, except one expert, were combined into a single standard observer group. The selection criteria for the observers followed the guidelines specified in ITU-R BT.500-14 [38], which ensured that none of the observers had any personal involvement in the design of the experiment. A Snellen chart test was also used to ensure that all observers had a normal or corrected-to-normal vision during the selection process.

### 3.4.2 Treatments

For our study, we selected a set of 20 images, consisting of four text-based images and eight each from fragments and natural scenes. We scanned these images using three devices and utilized CS for image enhancement. For CS, we adjust the minimum and maximum intensity values to encompass the full range of possible values. In the context of an 8-bit grayscale image, this means that the lowest intensity value in the image is expanded to the minimum possible intensity value of 0, while the highest intensity value is extended to reach the maximum value of 255. This stretching process is used to increase the dynamic range of the gray levels in the image. Equation (1) represents the general equation for contrast stretching. Data preprocessing and result analysis were performed using the open-source Python programming language.

$$OutputPixel = \left( \frac{InputPixel - MinInput}{MaxInput - MinInput} \right) \times 255 \tag{1}$$

Here, OutputPixel represents the resulting pixel value after contrast stretching, InputPixel is the original pixel value, MinInput and MaxInput are the minimum and maximum intensity values in the image, respectively. Next, we divided the images into 20 sets, each containing six images, three scanned images from each of the three devices and three enhanced versions of these images. Consequently, each set of images produced 15 pairs, resulting in 300 pairs shown to each observer in a randomized order. Typically, psychophysical experiments should not take more than one hour of total time for the observers. In our experiment, observers were not given a specific

time limit, but on average, we found that for 300 pairs, the median time taken by each observer was 30 minutes.

### 3.4.3 Instructions

At the beginning of the experiment, observers received instructions on the experiment. They were asked to compare two images displayed side by side and choose the one with better image quality. Observers were informed that the subjective test was being conducted to evaluate the performance of an alternative imaging approach for microfiche. Although observers were not provided with a specific definition of "quality" at the beginning of the experiment, we conducted a survey at the end to gather their feedback. This survey aimed to assess the observer's preferences concerning various predefined image quality attributes that they considered while rating each type of microfiche.

### 3.4.4 Viewing Condition

The experiment followed the guidelines outlined in ITU-R BT.500-14 [38] regarding viewing conditions on displays (ISO 3664). To ensure consistency, the monitor was calibrated using an Eye-one device prior to the experiment. The chromaticity of the white displayed on the color monitor was set to CIE standard illuminant D65, and the white's luminance level was set to 80 cd/m$^2$. The observers were seated approximately 80 cm away from the monitor, and the lighting in the room was dimmed to approximately 17 lux.

### 3.4.5 Subjective Data Processing

In a pair comparison experiment, the results can be presented as a winning frequency matrix that illustrates the relative frequencies with which each stimulus is preferred over the others. For example, in this study, comparing reproductions of three different devices (FSL, FBS, and IFS), participants are presented with all possible combinations of devices and asked to select their preferred option. The resulting data is recorded in a 3×3 raw data matrix. By aggregating the responses of all participants, a 3×3 raw frequency matrix is generated, providing a summary of the overall preferences for each device. This matrix visually represents the test outcomes, making it easier to interpret and draw some initial conclusions from the experiment data.

Using the frequency matrix, we computed the Binomial Sign test to show the statistical significance of the result obtained. To account for the possibility of type I errors resulting from multiple condition testing, we applied the Bonferroni correction [39].

Additionally, we computed the Z-score [40] based on Thurstone's law of comparative judgment [41]. It is a statistical measure that indicates how far a particular observation is from the mean in terms of standard deviation and allows us to draw statistical inferences about the differences between the two items being compared.

The most common approach to compute the 95% confidence intervals (CI) for Z-scores involves using the standard deviation ($\sigma$) and the number of observations (N) in the analysis and is carried out using Equation (2). When the 95% confidence intervals of the Z-scores do not overlap, we can conclude with 95% confidence that there is a significant difference.

$$CI = 1.96\frac{\sigma}{\sqrt{N}} \tag{2}$$

# 4 Results and Discussion

The results of this study involving both expert and rest of the observers for all images are shown in Figure 2. The Z-scores were calculated by comparing non-enhanced and enhanced images (using contrast stretching) from all three devices against each other. From analysis of Figure 2, it becomes apparent that the observers preferred IFS (both NE and CS) over FSL (NE and CS) and FBS (NE and CS), of which FSL and FBS had comparable responses for NE and CS respectively. In terms of observers' preference for enhanced versus non-enhanced images, the enhanced IFS images are most preferred over the non-enhanced IFS images. Non-enhanced images from IFS are preferred over the enhanced versions from the other two devices, making IFS the preferred device. However, images from FSL and FBS devices have comparable preference levels after enhancement, as shown by overlapping confidence intervals. Non-enhanced FSL images are preferred over non-enhanced FBS. It is important to note that this is for the current settings, devices, and images used for the experiment.

We also performed a separate analysis for the expert and the rest of the observers. The results of the experts were similar to those obtained from all combined observers. When analyzing the results for an expert, we found that IFS remained the preferred option. The Z-scores computed separately for the expert and all other observers are provided as a supplementary file (see Figure S3). Table 3 illustrates the result of hypothesis testing using a sign test for all images between the devices, indicating significant differences between the three devices. We also performed a sign test on the data from the expert, which produced similar results, except for the comparison

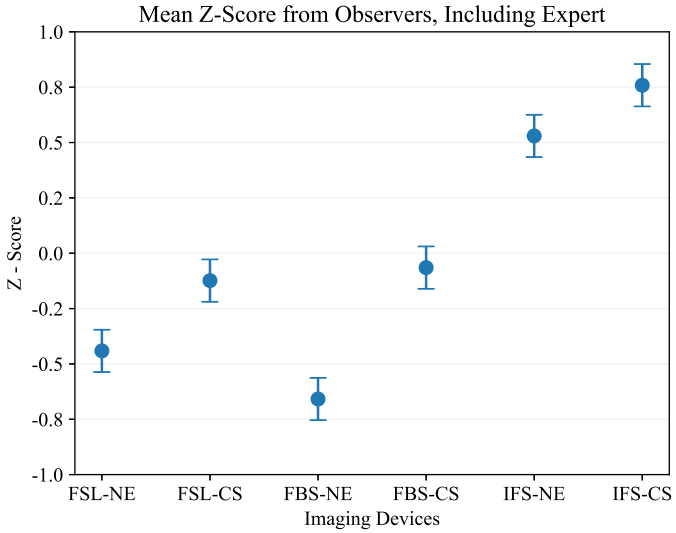



Figure 2: Comparison of Z-Score for all non-enhanced (NE) and contrast-stretched (CS) enhanced images, evaluated by all observers (including experts and rest of the observers), on all three imaging devices: flatbed scanner (FBS), in-house film scanning (IFS), and microform reader (FSL).

between FSL and FBS. The formulated null hypothesis that claimed that both devices produce similar results was true. Table S1 in the supplementary file contains a matrix that compares the significant test results for the expert.

In this experiment, three types of microfiche were used: text, natural scenes, and fragments. We also analyze user preferences for each microfiche category by comparing non-enhanced and enhanced versions of images from the enhanced and non-enhanced images from the three devices against each other. Both non-enhanced and enhanced versions from IFS are preferred for microfiche containing text (Figure 3a) and natural scenes (Figure 3b). However, for microfiche with fragments (Figure 3c), the enhanced versions from FSL and FBS have similar preferences to the non-enhanced version from IFS. However, the enhanced version from IFS remains the most preferred compared to all others. A non-enhanced version of the device IFS is always preferred against the other two devices in both enhanced and non-enhanced cases, indicating that the IFS device exhibits a higher capability for digitizing microfiche than the other two

Table 3: P-value for each pair of stimuli being compared; here each cell of a lower triangular matrix represents a pair of stimuli being compared; green cells indicate that the corresponding stimulus pair has a statistically significant difference.

| | **FSL** | **FBS** | **IFS** |
|---|---|---|---|
| **FSL** | | | |
| **FBS** | **2.78E-04** | | |
| **IFS** | **1.36E-126** | **1.40E-119** | |

devices. When comparing the performance between the FSL and FBS devices, it becomes difficult to determine which one exhibits better performance. For text-based microfiche, the non-enhanced image from the FSL device is equally preferred as the enhanced version from the FBS device, indicating that the FSL device may be considered better in this aspect. However, the enhanced version of the FBS device is preferred when evaluating natural scenes. Similarly, for microfiche with fragments, the non-enhanced images from the FBS device are less preferred, but both devices demonstrate similar preferences when considering the enhanced versions.

In addition, non-enhanced images from FBS remain the least preferred one for microfiche with text and fragments whereas for natural scene, FBS have similar preference as FSL. When comparing non-enhanced and enhanced versions, there is little or no difference in the case of microfiche with text, as indicated by the overlapping confidence intervals. The same applies to natural scenes, especially for IFS. However, there is a noticeable difference between the non-enhanced and enhanced versions in all other cases. A notable observation from the results is that the observers' ability to differentiate between the devices is becoming less distinct as we compare microfiche with text and natural scene; this distinction is even less for microfiche with the fragment. This can be seen in the range of Z-score values in Figure 3. Compared to natural scenes and fragments, observers took less time on average to make a decision regarding microfiche with text. It is possible that the devices used in this study were better able to distinguish their performance for simpler and more structured information, while their performance was less distinct for more complex and less structured data.

Table 4 displays the outcomes of hypothesis testing conducted on microfiche with text using various imaging devices, with and without enhancement. It indicates that IFS differs significantly from FSL and FBS, while there is no significant difference between FSL and FBS. It is also evident that when comparing enhanced and non-enhanced versions from the same imaging devices, the results indicate no statistically significant difference in the images for all three microfiche categories. Similar observations were

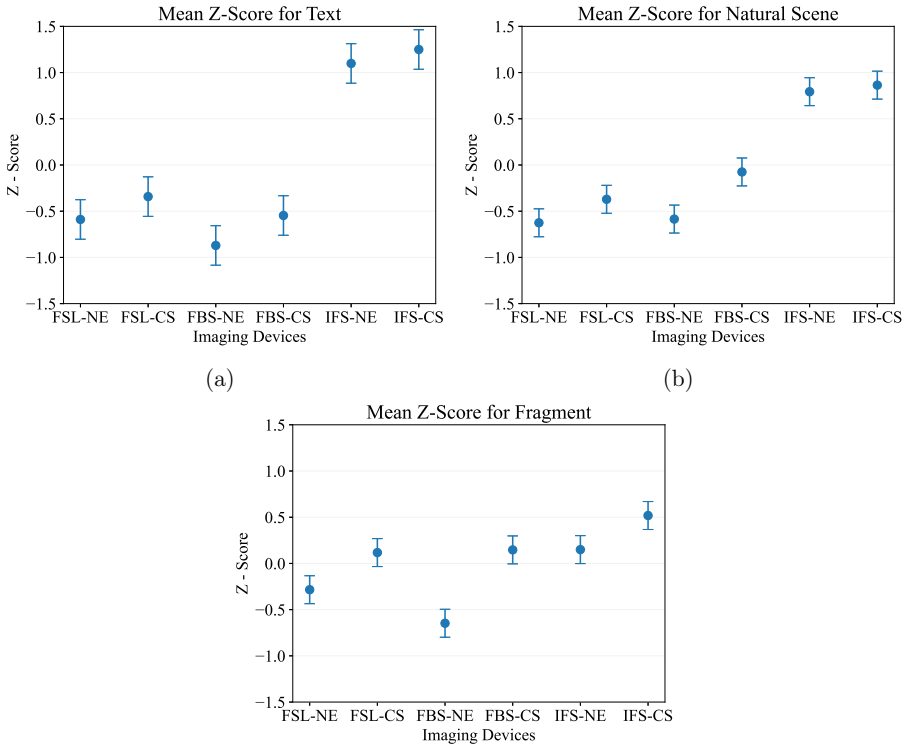



Figure 3: Z-scores at 95% confidence intervals for each microfiche category (i.e., with (a) Text, (b) Natural Scene, and (c) Fragment ) obtained by comparing non-enhanced (NE) and enhanced versions(CS) of images from all three devices, i.e., flatbed scanner (FBS), in-house film scanning (IFS), and microform reader (FSL), against each other.

made for the other two microfiche categories, namely natural scene and fragment, and the results of these categories are included in the supplementary file (Tables S2 and S3).

This study also examined observers' general preferences for enhanced versus non-enhanced images for each device, regardless of the type of microfiche being used. The results shown in Figure 4 demonstrate that observers consistently preferred enhanced

Table 4: The p-value for each pair of stimuli compared for microfiche with text is shown in a lower triangular matrix where each cell represents a pair of stimuli being compared. Green cells indicate a statistically significant difference between the corresponding stimulus pairs, while red cells indicate no statistically significant difference.

| | | FSL | | FBS | | IFS | |
|---|---|---|---|---|---|---|---|
| | | NE | CS | NE | CS | NE | CS |
| FSL | NE | | | | | | |
| | CS | 0.59151 | | | | | |
| FBS | NE | 0.00056 | 0.00443 | | | | |
| | CS | 0.0058 | 0.03209 | 0.54429 | | | |
| IFS | NE | 3.2E-09 | 8.4E-11 | 1.9E-19 | 3.1E-17 | | |
| | CS | 2.3E-09 | 5.9E-11 | 1.2E-19 | 2E-17 | 1 | |

images over non-enhanced images on all imaging devices. This indicates that the application of enhancement techniques improves the overall visual quality of the images, regardless of the specific type of microfiche or the imaging device used.

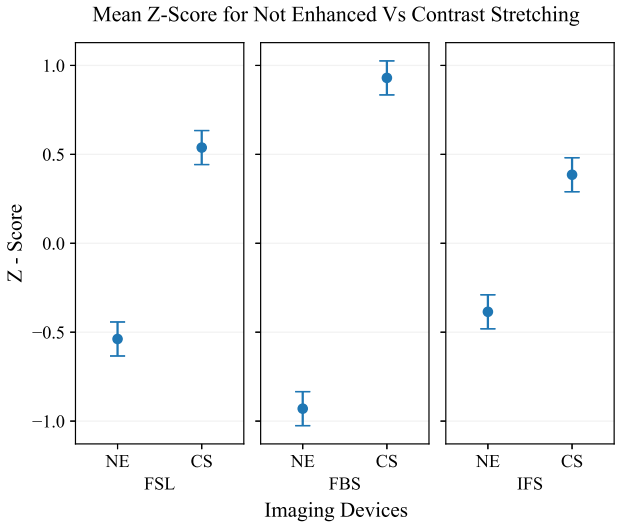



Figure 4: Z-scores, at 95% confidence intervals, between non-enhanced and contrast-stretched images across three imaging devices.

Table 5 displays the results of hypothesis testing conducted for enhanced versus non-enhanced images for all three imaging devices. Most of the cells are green, indicating a significant difference between the stimuli. However, results were not significantly different when enhanced versus enhanced compared in-between FSL, FBS, and IFS. This is not the case for non-enhanced images except for comparing IFS and FSL. To facilitate visualization, we have included the images in the supplementary file (Figure S5) that contain enhanced and non-enhanced images, one from each device.

Table 5: The p-value for each pair of stimuli compared for enhanced versus non-enhanced across three imaging devices; in a lower triangular matrix, green cells indicate a statistically significant difference between the corresponding stimulus pairs, while red cells indicate no statistically significant difference.

| | | **FSL** | | **FBS** | | **IFS** | |
|---|---|---|---|---|---|---|---|
| | | **NE** | **CS** | **NE** | **CS** | **NE** | **CS** |
| **FSL** | **NE** | | | | | | |
| | **CS** | **7.19E-17** | | | | | |
| **FBS** | **NE** | **0.000497** | **1.49E-30** | | | | |
| | **CS** | **2.11E-24** | **0.053128** | **6.43E-40** | | | |
| **IFS** | **NE** | **0.181417** | **2.04E-12** | **1.28E-06** | **4.8E-19** | | |
| | **CS** | **1.1E-13** | **0.356378** | **2.18E-26** | **0.0038** | **1.1E-09** | |

The observers were presented with one image from each type of microfiche, namely text, natural scenes, and fragments, and were provided with five options, including an option for user input. They were then asked to select which attributes they considered when rating the quality from these options. Figure 5 illustrates the result of the survey. The result infers that, when assessing the quality of microfiche, legibility is a crucial factor that all observers consider, particularly when examining text. Additionally, sharpness is another key attribute that observers prioritize when evaluating text and fragments on the microfiche. For natural scenes on the microfiche, contrast is the most significant attribute that observers consider when assessing quality. However, the lack of artifacts, such as noise, is regarded as a less important attribute across all three categories of microfiche. This means that the presence of minor imperfections, such as graininess or distortion in the images, would not significantly affect the overall quality assessment of the microfiche.

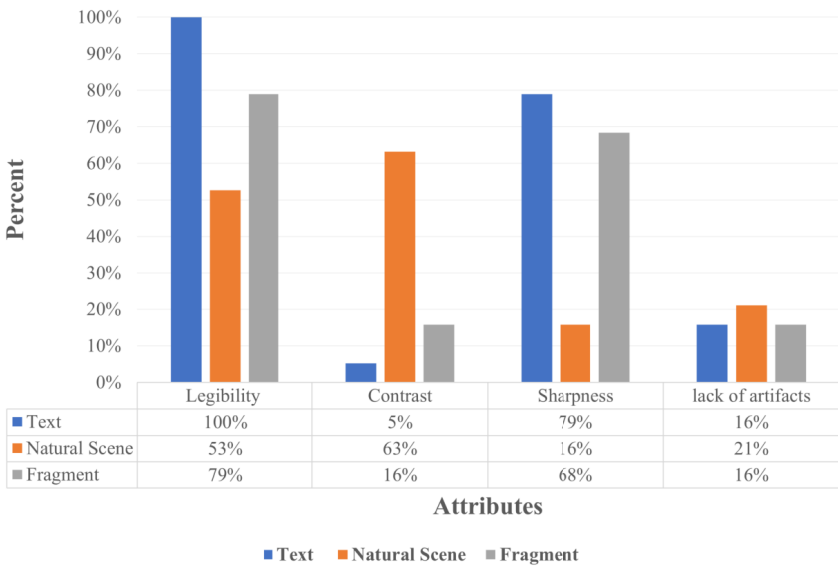



Figure 5: Results of a survey showing observer's preferences concerning various predefined image quality attributes, i.e., legibility, contrast, sharpness, and lack of artifacts that they considered most significant while rating each type of microfiche, namely text, natural scenes, and fragments.

# 5 Conclusion

Microfiche has been widely used to preserve historical documents, leading to the formation of valuable artifacts and essential elements of our cultural heritage. Despite its durability, microfiche can still suffer damage and require digitization for preservation and accessibility. However, traditional microfiche readers are not always available and are primarily designed for reading rather than digitization. This means that a digitization effort using a microfiche reader would also be significantly slower and time-consuming, consult Max. scan area and Max. fiche per scan in Table 1.

In this study, we evaluate the performance of alternative imaging devices compared to a traditional microfiche reader and the impact of image enhancement on image quality using subjective image quality assessment. The experiment was carried out

in a controlled environment with twenty-one participants, including an expert. The results of this study show that the reproduction of alternative devices was preferred over a traditional microfiche reader, indicating that alternative imaging devices can be a viable option for digitizing microfiche and improving access to historical collections. Additionally, results also demonstrate that the image enhancement techniques significantly improved image quality for all three categories of microfiche. Therefore, the adoption of alternative imaging devices along with image enhancement techniques could be a feasible approach to digitizing microfiche and facilitate better access to historical collections. This initiative could help conserve precious cultural heritage artifacts that exist only in microfiche format and enable more people to access them.

In future research, it would be beneficial to conduct objective image quality assessments for different quality attributes of microfiche, including contrast, sharpness, etc. This could involve exploring various existing image quality metrics and determining their relevance for assessing microfiche quality. Based on the results, new objective quality metrics could be developed or linked to provide more comprehensive and time-efficient evaluations of microfiche image quality. In addition, when digitizing microfiche technical drawings, it is crucial to ensure that the digitized versions faithfully capture the precise details and accuracy of the original drawings. Thus, further comprehensive studies can be conducted to investigate the geometric quality attributes, such as distortion, scale accuracy, resolution, etc., for microfiche digitization.

## Acknowledgment

The authors would like to thank Torleif Elgvin and Gregory High for providing access to the various materials and devices and Gjøvik Biblioteket for allowing us access to the Microform scanner needed for the formulation and execution of this study. This work is supported by The Lying Pen of Scribes–Manuscript Forgeries, Digital Imaging, and Critical Provenance Research funded by the Research Council of Norway (project nr. 275293) and CHANGE-ITN project funded by EU Horizon 2020(Marie Sklodowska–Curie, grant agreement No. 813789).

## References

[1] W De Haas. The microfiche. *Journal of the American Society for Information Science*, 9(2):99, 1958.

[2] Cristela Garcia-Spitz. Patrolling the past: Bringing the papua new guinea colonial-era reports into the digital realm. *UC San Diego: Library*, 2017.

[3] James M. Reilly. The IPI Storage Guide for Acetate Film. 1993, Retrieved from https://s3.cad.rit.edu/ipi-assets/publications/acetate_guide.pdf. Accessed: 27 June 2023.

[4] Kitti Canepi, Becky Ryder, Michelle Sitko, and Catherine Weng. Microform Terminology. In *Managing microforms in the Digital age*, pages 267–274. Association for Library Collections & Technical Services (ALCTS), 2013.

[5] South Carolina Department of Archives and History, Archives and Records Management Division. Microfilm and microforms. 2012, Retrieved from https://dc.statelibrary.sc.gov/handle/10827/6890. Accessed: 27 June 2023.

[6] ISO 18901:2010. Imaging materials – processed silver-gelatin-type black-and-white films – specifications for stability, 2010. Accessed on June 27, 2023.

[7] ANSI/IS0 8225-1995, ANSI/NAPM IT9.5-1996. Imaging materials – ammonia-processed diazo photographic film – specifications for stability, 1996. Accessed on June 27, 2023.

[8] ISO 18912:2002. Imaging materials – processed vesicular photographic film – specifications for stability, 2002. Accessed on June 27, 2023.

[9] Northeast Document Conservation Center (NEDCC). 6.1 Microfilm and Microfiche, 2017.

[10] Eibert Tigchelaar. On the unidentified fragments of "DJD" XXXIII and PAM 43.680: A new manuscript of "4QNarrative and poetic composition", and fragments of "4Q13", "4Q269", "4Q525" and "4QSb"(?). *Revue de Qumrân*, 21(3 (83)):477–485, 2004.

[11] Peter B Hirtle. The impact of digitization on special collections in libraries. *Libraries & Culture*, pages 42–52, 2002.

[12] William Saffady et al. *Computer-output microfilm*. American Library Association, 1978.

[13] Marius Pedersen, Jon Yngve Hardeberg, et al. Full-reference image quality metrics: Classification and evaluation. *Foundations and Trends® in Computer Graphics and Vision*, 7(1):1–80, 2012.

[14] Ping Zhao and Marius Pedersen. Extending subjective experiments for image quality assessment with baseline adjustments. In *Image Quality and System Performance XII*, volume 9396, pages 244–256. SPIE, 2015.

[15] Thomas Lee, Miaki Ishii, and Paul Okubo. Assessing the fidelity of seismic records from microfilm and paper media. *Seismological Society of America*, 93(6):3444–3453, 2022.

[16] Wendy M Duff and Joan M Cherry. Use of historical documents in a digital world: comparisons with original materials and microfiche. *Information Research*, 6(1):6–1, 2000.

[17] Syed Saqib Bukhari, Ahmad Kadi, Mohammad Ayman Jouneh, Fahim Mahmood Mir, and Andreas Dengel. ANYOCR: An open-source OCR system for historical archives. In *2017 14th IAPR international conference on document analysis and recognition (ICDAR)*, volume 1, pages 305–310. IEEE, 2017.

[18] Rose Holley. How good can it get? analysing and improving ocr accuracy in large scale historic newspaper digitisation programs. *D-Lib Magazine*, 15(3/4), 2009.

[19] Alee Rizwan Ali and Whitney Jackson Spivey. The archives of the future. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2021.

[20] Hilda Deborah and Dipendra J. Mandal. Evaluation of text legibility in alternative imaging approaches to microfiche digitization. In *Proc. IS&T Archiving*, pages 96–101, 2021.

[21] Zhan Su, Byung-Ryul Ahn, Ki-Yol Eom, Min-Koo Kang, Jin-Pyung Kim, and Moon-Kyun Kim. Plagiarism detection using the levenshtein distance and smith-waterman algorithm. In *2008 3rd International Conference on Innovative Computing Information and Control*, pages 569–569. IEEE, 2008.

[22] Gursharn Singh and Anand Mittal. Various image enhancement techniques-a critical review. *International Journal of Innovation and Scientific Research*, 10(2):267–274, 2014.

[23] Marius Pedersen, Nicolas Bonnier, Jon Yngve Hardeberg, and Fritz Albregtsen. Attributes of image quality for color prints. *Journal of Electronic Imaging*, 19(1):011016–011016, 2010.

[24] Ke Gu, Guangtao Zhai, Xiaokang Yang, Wenjun Zhang, and Min Liu. Subjective and objective quality assessment for images with contrast change. In *2013 IEEE International Conference on Image Processing*, pages 383–387, 2013.

[25] Ke Gu, Guangtao Zhai, Weisi Lin, and Min Liu. The analysis of image contrast: From quality assessment to automatic enhancement. *IEEE Transactions on Cybernetics*, 46(1):284–297, 2016.

[26] Xiaolin Wu. A linear programming approach for optimal contrast-tone mapping. *IEEE transactions on image processing*, 20(5):1262–1272, 2010.

[27] Ching-Chung Yang. Image enhancement by modified contrast-stretching manipulation. *Optics & Laser Technology*, 38(3):196–201, 2006.

[28] Giorgio Trumpy, Jon Y. Hardeberg, Sony George, Barbara Flueckiger, Barbara Cattaneo, Marcello Picollo, Filippo Cherubini, and Veronica Marchiafava. Multispectral capture of film colors with led. In *Colour photography and film: Sharing knowledge of analysis, preservation, conservation, migration of analogue and digital materials*, pages 195–198. Gruppo del Colore–Associazione Italiana Colore, 2021.

[29] William R Hawken. *Evaluating Microfiche Readers: A Handbook for Librarians*. Washington: Council on Library Resources, 1975.

[30] George J. Brooke and Helen K. Bond. *Allegro Qumran Photograph Collection*. Brill, Leiden, The Netherlands, 1999.

[31] Emanuel Tov and Stephen Pfann. *Companion Volume to the Dead Sea Scrolls on Microfiche Edition: Published under the Auspices of the Israel Antiquities Authority*. Brill, Leiden, The Netherlands, 1995.

[32] Rafał K Mantiuk, Anna Tomaszewska, and Radosław Mantiuk. Comparison of four subjective methods for image quality assessment. *Computer Graphics Forum*, 31(8):2478–2491, 2012.

[33] Margaret H Pinson and Stephen Wolf. Comparing subjective video quality testing methodologies. In *Visual Communications and Image Processing 2003*, volume 5150, pages 573–582. SPIE, 2003.

[34] Brian W. Keelan. *Handbook of image quality: characterization and prediction*. CRC Press, 2002.

[35] Marius Pedersen. *Image quality metrics for the evaluation of printing workflows*. PhD thesis, University of Oslo, 2011.

[36] Herbert Aron David. *The method of paired comparisons*, volume 12. London, 1963.

[37] Khai Van Ngo, Jehans J. Storvik, Christopher A. Dokkeberg, Ivar Farup, and Marius Pedersen. QuickEval: A web application for psychometric scaling experiments. In Mohamed-Chaker Larabi and Sophie Triantaphillidou, editors, *Image Quality and System Performance XII*, page 93960O. SPIE, 2015.

[38] BT.500-11 Recomendation ITU-R. Methodology for the subjective assessment of the quality of television images. *International Telecommunication Union*, 2019.

[39] J Martin Bland and Douglas G Altman. Multiple significance tests: the bonferroni method. *British Medical Journal(BMJ)*, 310(6973):170, 1995.

[40] Peter G Engeldrum. Psychometric scaling: avoiding the pitfalls and hazards. In *Image Processing, Image Quality, Image Capture Systems Conference (PICS)*, pages 101–107, 2001.

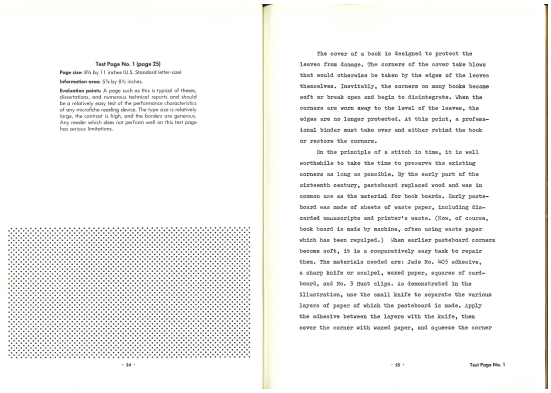
[41] Louis L Thurstone. A law of comparative judgment. In *Scaling*, pages 81–92. Routledge, 2017.

[42] Zeutschel GmbH. Microfilm scanner delta+ Advanced Microfilm Scanner. Retrieved from https://www.zeutschel.de/en/produkte/deltaadvanced/. Accessed: 16 May 2023.

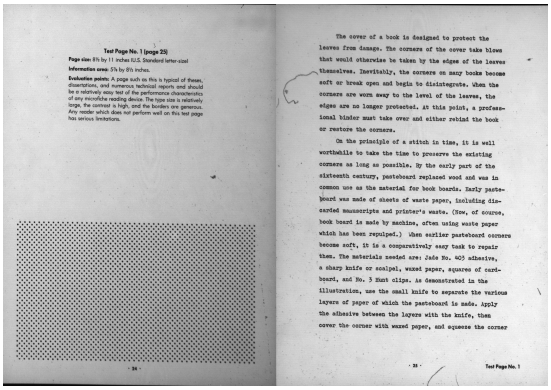
[43] EPSON. Epson Perfection 4870 Photo. Retrieved from https://www.epson.de/de_DE/produkte/scanner/consumer/epson-perfection-4870-photo/p/327. Accessed: 28 June 2023.

## Supplementary Materials

### S I. Original document image and its magnified image from the microfiche.




(a)




(b)

Figure S1: Image showing the original document image (a) and its magnified image from the microfiche(b). We can observe various artifacts within the microfiche image (b), although distinguishing between artifacts originating from dust, dirt, or the microfiche itself can be challenging. Due to the magnification process, these artifacts become prominently visible in the images.

## S II. Microfiche readers used in this study.

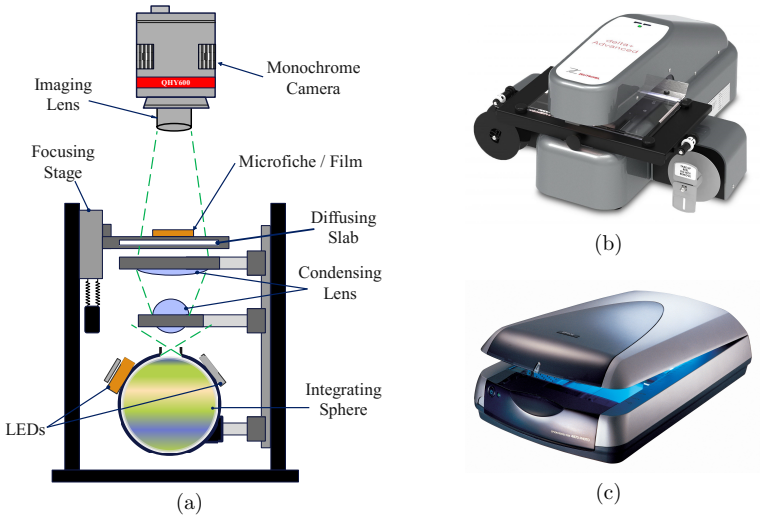



Figure S2: Microfiche readers (a) schematic of the in-house film scanner, *image adapted from* [28], (b) microform reader, *image Source* [42], and (c) flatbed scanner, *image Source* [43].

## S III. Z-Score for comparing enhanced and non-enhanced images.

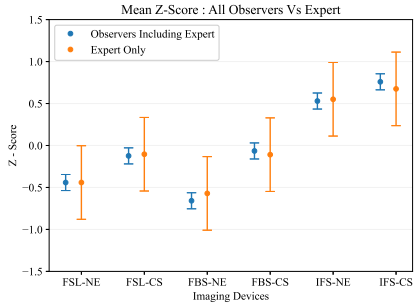



Figure S3: Comparison between the Z-Score from all observers and an expert, with 95% confidence intervals, for non-enhanced and contrast stretching across all three imaging devices.

## S IV. Result of the sign test showing the expert preference between devices.

Table S1: P-value for each pair of stimuli being compared; here each cell of a lower triangular matrix represents a pair of stimuli being compared; Green cells indicate a statistically significant difference between the corresponding stimulus pairs, while red cells indicate no statistically significant difference.

| | FSL | FBS | IFS |
|---|---|---|---|
| **FSL** | | | |
| **FBS** | 0.91098 | | |
| **IFS** | 1.3E-05 | 1.48E-07 | |

## S V. Result of the sign test comparing enhanced and non-enhanced images of all three devices for microfiche with a natural scene.

Table S2: The p-value for each pair of stimuli compared for microfiche with natural scene is shown in a lower triangular matrix where each cell represents a pair of stimuli being compared. Green cells indicate a statistically significant difference between the corresponding stimulus pairs, while red cells indicate no statistically significant difference.

| | | FSL | | FBS | | IFS | |
|---|---|---|---|---|---|---|---|
| | | NE | CS | NE | CS | NE | CS |
| **FSL** | **NE** | | | | | | |
| | **CS** | 0.72629 | | | | | |
| **FBS** | **NE** | 0.891 | 0.57818 | | | | |
| | **CS** | 0.14951 | 0.06311 | 0.21616 | | | |
| **IFS** | **NE** | 1.5E-20 | 4E-22 | 8.1E-20 | 2.5E-15 | | |
| | **CS** | 2.2E-18 | 7E-20 | 1.1E-17 | 1.9E-13 | 0.58543 | |

## S VI. Result of the sign test comparing enhanced and non-enhanced images of all three devices for microfiche with a fragment.

Table S3: The p-value for each pair of stimuli compared for microfiche with a fragment is shown in a lower triangular matrix where each cell represents a pair of stimuli being compared. Green cells indicate a statistically significant difference between the corresponding stimulus pairs, while red cells indicate no statistically significant difference.

| | | FSL | | FBS | | IFS | |
|---|---|---|---|---|---|---|---|
| | | NE | CS | NE | CS | NE | CS |
| FSL | NE | | | | | | |
| | CS | 0.90822 | | | | | |
| FBS | NE | 0.05191 | 0.07648 | | | | |
| | CS | 0.34494 | 0.44047 | 0.34661 | | | |
| IFS | NE | 4.6E-05 | 2.2E-05 | 1.7E-09 | 4.3E-07 | | |
| | CS | 0.00051 | 0.00027 | 5.6E-08 | 8.3E-06 | 0.57288 | |

## S VII. Images of the microfiche with fragments showing before and after enhancement.

(a) FBS_NE (b) FBS_CS

(c) FSL_NE (d) FSL_CS

(e) IFS_NE (f) IFS_CS

Figure S4: Microfiche images with fragments showing images with enhancement (_CS) and without enhancement (_NE). Image from flatbed scanner (FBS), microform reader (FSL), and in-house film scanner (IFS).

ISBN 978-82-326-7412-1 (printed ver.)
ISBN 978-82-326-7411-4 (electronic ver.)
ISSN 1503-8181 (printed ver.)
ISSN 2703-8084 (online ver.)





NTNU

Norwegian University of
Science and Technology